SELECT COMMITTEE TO INVESTIGATE THE
JANUARY 6TH ATTACK ON THE U.S. CAPITOL,
U.S. HOUSE OF REPRESENTATIVES,
WASHINGTON, D.C.

DEPOSITION OF:  J. JOHNSON

Wednesday, September 7, 2022

Washington, D.C.

The deposition in the above matter was held via Zoom, commencing at 12:08 p.m.
Present:  Representatives Aguilar and Raskin.
Appearances:

For the SELECT COMMITTEE TO INVESTIGATE
THE JANUARY 6TH ATTACK ON THE U.S. CAPITOL:

STAFF ASSOCIATE
SENIOR ADMINISTRATIVE ASSISTANT
PROFESSIONAL STAFF MEMBER
INVESTIGATIVE COUNSEL
CHIEF CLERK
COUNSEL
SENIOR COUNSEL & SENIOR ADVISOR

For WITNESS:

ALEXIS RONICKHER
With that, we can go on the record at 12:08 p.m.

Good afternoon.

This is the deposition of J. Johnson conducted by House Select Committee to Investigate the January 6th Attack on the United States Capitol pursuant to House Resolution 503. This will be a staff-led deposition, though members may, of course, choose to ask questions as well.

At this time I would like to note for the record that during this deposition, staff will refer to the deponent as J. Johnson in order to protect their identity. Any references to their legal name will be redacted from the record.

I want to ask the witness if you agree to that setup.

The Witness. I agree.

Thank you.

My name is [redacted]. I'm an investigative counsel with the select committee. With me today from the select committee staff are [redacted], senior counsel; [redacted], professional staff member; and [redacted], professional staff member -- senior administrative assistant. I'm sorry. We also have chief clerk [redacted]. And I will announce any additional staff members as they join the Zoom.

With us today we also have Representative Jamie Raskin, and if any other members join, we will announce them as well.

I'll begin the questioning today, although other staff will likely join in as well. If at any point for any staff questions you don't understand them, please ask us to repeat the question, and we'll do our best to rephrase.

J. Johnson, you have been subpoenaed by the select committee to compel your presence at this deposition. Can you confirm that you've received a subpoena from the
The Witness. I confirm.

Thank you.

And you understand that you're appearing here today pursuant to the deposition subpoena?

The Witness. I understand.

Thank you.

Because this is a deposition, you may only refuse to answer a question to preserve a privilege that's recognized by the select committee, such as the Fifth Amendment. If you refuse to answer a question based on a privilege, we may either proceed with the deposition, or seek a ruling from the chairman of the select committee on the objection. If the chairman overrules such an objection, then you'll have to answer the question.

Do you understand?

The Witness. I understand.

Thank you.

My goal today is to ask questions relevant to the select committee's investigation with the hope that you'll answer. If you have an objection or a privilege assertion, we'll ask that you or your counsel assert it for the record. Of course, we don't anticipate having that problem today. We appreciate your cooperation. But if it does happen, we may seek to clarify the purpose of this objection. Ultimately, the more detail you can provide in an objection, the easier it will be for the select committee and the chairman to evaluate your claim.

There is an official reporter transcribing the record of this interview. Please wait until each question is completed before you begin your response, and we'll try to wait until your responses are completed before we begin our next question.
The stenographer cannot easily record nonverbal responses, like shaking your head or giving a thumbs up, so it's important that you answer each question with an audible, verbal response, and that we try not to talk over each other. Of course, I don't anticipate this being a problem today either.

We ask that you provide complete answers based on your best recollection to each question. If a question is not clear, you can ask for clarification. If you don't know an answer, please just say so. Similarly, if at any point you need to discuss something with your attorney in private, please let us know, and we'll take a break so you can have that conversation.

I want to reiterate that this interview is under oath, and you're obligated under Federal law to tell the truth just as if you were speaking to the FBI or to DOJ. It is unlawful for you to deliberately provide false information to Congress. For this interview, providing false information could result in criminal penalties for perjury and/or false statements. And this is something that we tell all of our witnesses.

Do you understand?

The Witness. I understand.

Thank you.

You are not obligated to keep the fact of this deposition and what we discuss confidential. You're free to tell whomever you wish that you met with us, including the substance of what you discussed. But we will not share any details of what we discussed today unless it's authorized by the committee.

Okay. Do you understand?

The Witness. I understand.

All right. So now, because this deposition is under oath, would you please raise your right hand to be sworn in by the reporter.
The Reporter. Do you solemnly declare and affirm under the penalty of perjury that the testimony you are about to give will be truth, the whole truth, and nothing but the truth?

The Witness. I do.

EXAMINATION

Q So without objection, I would like to start with some general background questions about your time at Twitter.

So can you give us a sense of when you worked at Twitter?

A Yes. I joined Twitter in early 2020, and I was at the company until late 2021.

Q Thank you.

And during that time what teams did you work on?

A I was on the safety policy team.

Q Great. Thank you.

And can you give us a general sense of your responsibilities while you were a member of the safety policy team?

A Yes. My role was to create policy for content on Twitter, so creating the rules for what you can and cannot say, as well as dealing with high-profile escalations. So if a user potentially violated our rules, it was my job to understand whether or not it was, indeed, a violation and subsequently what remediation, whether they should be suspended permanently or for a handful of hours.

Q Understood.

So during that time frame, I would like to ask about the nature of your
involvement with Twitter’s preparation for and response to the January 6th attack on the Capitol. It seems like you would have had some involvement given how you just described your responsibilities.

A Yes, that’s right.

Q So can you give us a sense of the nature of your involvement with the preparation for the January 6th attack and the response to it?

A Yes. So our team, as mentioned, we wrote the rules for the platform, and so, I was also aware of preparations that we did ahead of the election and post-election. So within that there was a series of exercises we went through, policy developments, and coordinating our response to events, including January 6th.

Q And did that deal specifically with how to handle accounts that were posting potentially violent content on the platform?

A Yes, yes, that did. We would evaluate certain risks and what policies that needed to be deployed in order to respond to those risks.

Q Thank you.

And we will get into some of those policies and proposed policies shortly.

And did you have a role -- can you describe to us the nature of your role in the suspension of President Trump’s account after January 6th?

A Yes. My role was to evaluate Donald Trump’s tweets and understand whether or not they violated our policies.

So on January 6th, Twitter made the decision that if Trump posted another tweet that violated our policies, he would be permanently suspended from the platform. So anytime Donald Trump tweeted after January 6th, my team would evaluate that to understand whether it violated our policies. And I was a part of the team that would assess that tweet and, ultimately, was a part of the assessment that ended up in his
permanent suspension from Twitter.

Q Thank you.

And we'll certainly get back to that later on in the deposition as well. I appreciate the background.

And you said you left Twitter in 2021?

A That's right.

Q Can you give us a general sense of what prompted you to leave Twitter?

A There was a lot of frustrations about the lack of leadership and the fact that Twitter was not responding to the concerns that I and others on my team had repeatedly raised, and I did not see a future for myself at the company where I would be able to address some of those risks in a way that I felt comfortable with, which ultimately left to my departing the company.

Q Got it.

So would you say that Twitter's response to the January 6th attack played a role in your decision to leave the company later that year?

A Yes, it absolutely did.

Q Could you expand upon that at all?

A Yes. I had asked the company, my team, leadership on my team, multiple times to conduct the retrospective for what happened leading up to the attack on the Capitol and the suspension of Donald Trump, and that was never granted by leadership.

I was told that it wasn't a priority for the company. And that was concerning to me, in addition to the fact that there were other elections coming up globally, be it in Brazil later this year, or the U.S. midterms. And at that time, even though it was 2021 and that was far in the future, I had some serious concerns about the direction the company was headed, especially given there was a reluctance to address any of the concerns that I had
repeatedly raised.

Q Thank you. And we'll return to some of those questions, particularly about retrospective, towards the end of our conversation as well. But I appreciate the general background that you just gave, and I want to move into some more granular discussion of what was happening at Twitter leading up to the election.

But before I do, I wanted to pause and see if Representative Raskin had any questions at this point.

Seeing none, we can continue.

Mr. Raskin. Actually, forgive me.

Of course.

Mr. Raskin. Yeah. I wonder if you could just expand on the last response by explaining how the company's response to the events of January 6th ranked among other causes in your decision to leave the company. Would you say that was the major reason that you left Twitter?

The Witness. Yes, that was one of the major reasons. I would say it would rank one.

Mr. Raskin. And the other reasons were consistent with your disappointment about Twitter's response to the events of January 6th?

The Witness. That's correct, yes.

Mr. Raskin. Okay. Thanks much.

I yield back.

Thank you, Mr. Raskin.

And I'll pause periodically to see if Mr. Raskin or anyone else has questions.
Q: So at this point I would like to move into the time period leading up to Election Day 2020. And you were on the safety policy team at that point, correct?

A: Correct, I was.

Q: So, in your view, what was the state of the safety policy team in the summer and fall of 2020? Did you think it was adequately resourced?

A: I did not think it was adequately resourced. The team was very disorganized. We had multiple members within the team leave the company, and so, there was not any sort of organization or clear understanding of how safety policy, who owns 20-plus policies that govern Twitter, was going to plan a response to the U.S. election.

Q: So when you said safety policy owned 20-plus policies, can you explain in laymen's terms what that means and what safety policy's Role was in sort of responding to election-related threats?

A: Yes. Safety policy owns -- what I mean by "own" is we are responsible for the development and implementation of over 20 policies, including abusive behavior, glorification of violence, privacy. The list goes on. So it would be our job to make sure that those policies are robust and adequately written and implemented to address the risks ahead of the U.S. election.

Q: So your team was responsible for reviewing instances of, for example, glorification of violence that was occurring on the platform, all across the platform in the months before the election?

A: That's correct, yes.

Q: And so, it was your view that your team was not big enough to handle this task or not organized enough?

A: A combination of the two. I would say that we were grossly understaffed
to address the risks leading up to the U.S. election, and the team was very new. My supervisor, who was leading our plans ahead of the U.S. election, had only been at the company a handful of months. I had only been at the company, you know, a handful of months before the election.

So a combination of the fact there was no organizational structure; there was not as much awareness of how teams interact with one another; in the event of a crisis, what are all of the remediations that we could use? There was a lack of clear processes for if we needed to approve a new policy development or if we needed to implement our policies to address the risks. All of that was very unclear at the time.

Q Thank you.

So can we zoom out for a moment and talk about the actual content that your team was seeing on the platform in those months prior to the election? What sort of trends were concerning to you?

A Yeah. The rhetoric on the platform we were seeing was increasingly hostile. We were getting multiple escalations a day, and escalations are more high-profile potential violations of our policy lines. And we were seeing that they were all more political in nature, which for a global company, it could pertain to anything, but we were seeing all escalations pertaining mostly to the United States and the upcoming election.

In the summer of 2020, it started with threats against Census workers, and it escalated from there. And we were concerned that our policies might not adequately address the risks we were seeing on the ground and the increasing tensions.

Q So when you say that your policies were not adequate to meet the threat on the ground, what particular shortcomings did your team identify in your policies that made you worry?
So there are two categories I would say that made our team worry. The first is the policy themselves. And a gap that we became aware of where there were coded incitement to violence and rhetoric that was more ambiguous, so it didn’t clearly violate our policy lines, but it was problematic content. So that’s one category where the policies themselves needed to be stronger.

And the second, as mentioned earlier, was the lack of clear processes, so who’s responsible for signing off on whether or not the tweet violates our rules? What time frame does that need to happen?

So I would say both of those issues were major concerns at the time.

Q So would it have been helpful for you at the time to have clear indications of where tweets needed to be approved, action against tweets that needed to be approved, and how long a turnaround time you had to review a tweet, that sort of thing was not something you were provided with?

A Yes, that’s correct.

Q And we’ll get back to the coded incitement issue shortly, and I actually have some documents to show you on that point. But one question that I had about the evolution of these trends that were more political and hostile in nature, sort of asking you to think back to the summer of 2020 and other things that were going on. Obviously, there were a lot of protests across the country for Black Lives Matter and some violent counter protests as well, and you also had protests that were armed related to COVID-19 restrictions.

So I’m wondering if your experience on the safety policy team exposed you to problematic content related to those issues and if those issues intercepted with election-related content?

A Yes, they definitely intersected in the sense that a lot of the protests over
the summer, you know, there was an undertow of political rhetoric associated with that. Donald Trump, the first time that we used the public interest interstitial was in response to one of those protests. So it was definitely connected, and I think that that should have been enough of a red flag to the company that we were seeing increased tensions across the country.

Q Got it.

So I want to back up there for a second, because you said something that I want to tease out.

You mentioned the first time that Twitter applied a public interest interstitial to President Trump's tweets was over the summer. Can you just explain for the record what a public interest interstitial is and give a sketch of the scenario where it was first applied?

A Yes. So the public interest interstitial, it was developed and approved before I joined the company. But what its intent is that rather than deleting a tweet that violates our policies, if it is in the public’s interest to see what that tweet is, we would put a label over it and reduce its distribution so people could still see what was said. And the only people where you would apply the public interest interstitial, or PII to, are accounts of elected officials; that being Donald Trump.

And it was joked amongst colleagues that this -- the development of the PII was in response to Donald Trump being elected President and a way to deal with his borderline violative tweets.

Q So essentially the idea is that Twitter was not able to take down President Trump's tweets, and this was a way to keep his tweets up but acknowledge their violative nature?

A That's correct, yes.
Q And so the first tweet, I believe, that the PII was applied to, the sort of warning message, was about the looting starting and the shooting starting?

A That’s correct. In response to the protests that were happening over the summer of 2020, Donald Trump tweeted along the lines of, "When the looting starts, the shooting starts." And the team -- I was on safety policy. We found that to be in violation of our policy against glorifying violence, and we applied the public interstitial to his account for the first time.

Q And if I’m remembering correctly, President Trump was half a year later -- or over a half a year later eventually suspended from Twitter for the violation of that same policy, glorification of violence, correct?

A That’s correct, yes.

Q So when the safety policy team was looking at the tweet, that initial tweet in the summer, did you see responses to that tweet or traffic elsewhere on the platform that made you concerned about the effect that tweet was having on the discussion on Twitter?

A Yes. So once we applied the public interest interstitial, we saw there was huge uproar, for lack of a better term, because of that action, and there were users on the platform who were copying and pasting that exact language and retweeting it, and that was actually a first instance where we saw our teams weren’t able to handle the response in the sense that because users were tweeting his violative tweet again, we didn’t have a plan in place to address that content and the response to that.

Mr. Raskin, do you have a question?

Q So essentially you saw folks elsewhere on the platform copying what the President was saying, and your team raised concerns that Twitter didn’t have a plan to
deal with the proliferation of this potentially insightful content?

A Yes, that's correct.

Q And do you recall any specific conversations where you talked with Twitter leadership about the need to make a plan to deal with this eventuality?

A My memory is not totally firmed up, but I do believe that we talked to leadership, but I also believe that there was actually a document that existed to try and piece together a plan of attack for what it would look like if folks kept retweeting the violative content that the President tweeted. But that was something that was more just for our team and was more documentation purposes rather than an organized plan in the event that this happened again in the future.

Q In your experience, have there ever been another instance of Twitter having to have these conversations about a world leader using his account to essentially encourage his supporters to tweet similar violative content?

A Not to my knowledge, no.

Q And was it your sense in the summer of 2020, before we got into the heart of the election season, that Twitter was reluctant to take this action, this initial action against President Trump's account?

A Yes, absolutely. It took a massive movement over the summer of 2020, and his response to that for us to apply what is, in my opinion, not a strict form of remediation on his tweet.

Q And why do you think Twitter was reluctant to do so?

A I think they were reluctant to action the President’s account because he was the President, and they were concerned about what that would mean for the company.

Q Understood. And so -- oh, Mr. Raskin.

Mr. Raskin. Have there been other cases of public officials who have received
the PII tag?

The Witness. Yes. Other public officials have received the public interest interstitial after we applied it to Donald Trump's account. The first time we used it, I believe, was for Donald Trump.

Mr. Raskin. I see. And -- well, I don't know if you were going to be getting into this, Jacob, but -- Jacob, are you planning to follow then the history of what follows after this original designation?

Mr. __________. Yes. And we're going to move back into the election period, and then we'll get into some of the other policies, too.

Mr. Raskin. Okay. Great.

I yield back.
Q So on Mr. Raskin’s general train of thought, I was curious to ask you about what you were seeing in response to President Trump’s tweets leading up to the election. Were there similar instances where his supporters were mimicking his claims about election fraud, mail-in voting, or other borderline violative content that he was propagating prior to the election?

A Yes, absolutely. I think that Donald Trump would tweet something, and the theme of whatever he was tweeting, be it election fraud, mail-in voting is not secure, you would see the ripple effects of that from his supporters and his base on Twitter.

Q And was your team, the Safety Policy Team, nervous at that point, September, October, 2020, that you were unable to deal with the potential harmful effects of that ripple given the tools that were, at that point, in your toolbox?

A I was concerned about that. I can’t speak to leadership, but I was certainly concerned. There was -- I was aware of an election threat model that detailed potential threats leading up to the election; for example, the President violating our policies, someone sharing President Biden’s private information, and we would rank what is the likelihood of that happening and then what is our preparedness for that. And the section for safety policy, I was aware of there was a lot of red on that spreadsheet, which indicated that we were not prepared and there was a high risk of that happening.

Q So what would have made you more prepared? And what was your team then -- you, yourself, what were you advocating for to make you more prepared?

A Yeah. There needed to be a clear understanding of the application of our policies in a consistent manner. At the time, decisions felt very arbitrary and dependent upon senior leadership and whether they were feeling external pressure, at least that’s
what it felt like for me and -- I'll just speak for myself. And then in addition to that, as mentioned earlier, clear processes to adequately understand who signs off on this, and what does that process look like. It was extremely unclear, both the application of our policies, and then how we actually get them off the platform.

Q So you mentioned that these decisions seemed arbitrary and based on whether Twitter was feeling external pressure. Would you be able to expand upon what you mean by external pressure?

A Yes. There were instances in which, over the course of my tenure at the company, we would see a high-profile tweet, but we would also get wind that maybe senior leadership was meeting with an interested party. So if it was a tweet from someone who was right-leaning, and then senior leadership was meeting with an organization that represented either that elected official or someone associated with them, they would -- that would feed into the decisions and the analysis of the tweet. So it felt as though when senior leadership had external pressure, be it from whomever they might be meeting with, whomever they might be talking to, that definitely fed into the decisions and the way that we analyzed the tweets. You know, a lot of these decisions can be argued one way or the other, and we were told -- and this leads to why it felt like the decisions were arbitrary. It often felt like we were arguing to leave certain things up when it didn’t feel right to do so because that was not how we had applied our policies for the rest of the Twitter community.

Q So how did this play out in the pre-election moment when, for example, there were a lot of claims of voter fraud emanating from President Trump's allies? Did that have an impact on how Twitter was analyzing how to action these tweets?

A One of the things that safety policy worked on and I was aware of was election guidance, which was an attempt to create clarity and create structure to the two
issues that I mentioned, the application of our policy, and the process associated with
that. This guidance was meant to remove that external pressure that I was talking about
earlier, and it was meant to try and make our policy decisions more consistent, more
defensible. This guidance was approved. However, to my knowledge, it was never
implemented. So even though it was technically shipped, no one ended up utilizing it in
the months leading up to the election.

I just want to note for the record that Representative Pete Aguilar has
joined as well. Welcome.

I'm not sure if you want to ask any questions at this point, Mr. Raskin or Mr.
Aguilar?

Mr. Raskin. I do have a question.

Were you involved in the preparation of that guidance that you just referred to?

The Witness. I was, yes.

Mr. Raskin. And was there anything in that guidance that addressed the
question of whether or not elected officials should be treated the same as everybody else
or whether their tweets should be regarded under a higher standard, or a more relaxed
standard, or are they supposed to just meld in with the rest of the population? Did you
ever address that problem?

The Witness. That was not addressed specifically in the guidance. However,
the guidance aims to create consistency. So if it was implemented and followed the way
that it was written and meant to be interpreted -- and, again, this is more conjecture
because it never was used -- then elected officials should be treated the same and we
should be applying the policies consistently, because that was the overarching goal of this
guidance was making sure that the policies are applied consistently, we can understand
when something is a violation as opposed to whatever leadership might be feeling that
day.

Mr. Raskin. When you referred to the seeming arbitrariness of the application of the different policies that are out there, did that arbitrariness relate just to arbitrary and selective treatment of different substance, not treating the substance in coherent or consistent ways, or did it also apply to treating different tweeters, different speakers in different ways?

The Witness. I want to make sure I'm understanding your question.

Mr. Raskin. Was it arbitrary just as to substance, or was it arbitrary also with respect to who the tweeter was?

The Witness. I understand. To both, I would say. It could depend upon who tweeted it. I would say that that definitely fed into the decision and what made something seem more arbitrary. So to be clear whether it was Donald Trump or an elected official, it felt as though we were not applying our policies consistently in those instances, and that, also, the substance of the tweet would also feed into the arbitrary nature of the decision. So it was both the substance of the tweet and who tweeted it.

Mr. Raskin. Okay. And if you could just give us some ballpark portrait of how many tweets would be taken down on a given day.

The Witness. I'm not entirely sure. My team was more the last resort to deal with these high-profile escalations. So I can answer on my team what that looked like leading up to the election. We definitely saw an uptick in escalation. So we might have seen 10 tweets come to us a day, but it was, I would say, a mixed bag of leaving it up or taking it down.

It's important to note our team was not the operational arm of Twitter, so we didn't have as much visibility into how many tweets were being taken down a day. We only saw kind of the very tip of the iceberg and the more borderline high-profile gray area
cases.

Mr. Raskin. Got you. Thank you.

I yield back.

Thank you, Mr. Raskin.

BY:

Q And while we're on the topic of how different accounts were perhaps treated differently, I want to ask if you're familiar with the very important tweeter policy that Twitter had to sort of give other teams a heads up if action was going to be taken against high-profile accounts?

A Yes, yes. Very important tweeter, or a VIT, so we would give our public policy team and our comms team and leadership a heads-up that we had determined a certain tweet was in violation of our policy, and then we would wait, I believe, if memory serves, somewhere around 4 to 6 hours before actually taking the tweet down. And there were -- if my memory serves, there were oftentimes where the public policy team would communicate to whomever violated the tweet, so they were aware that this was going to happen as well.

Q And in terms of how Donald Trump's account was tweeted -- excuse me -- was treated compared to other VIP accounts, do you have a sense of whether additional provisions were needed to take action against President Trump's account while he was in office, or whether there were special chains of permission to consider a PII, or a more dramatic action against his account?

A Anytime there was an escalation involving Donald Trump's account, there needed to be signoff from Jack Dorsey, and that was not the case for other VITs, and there were a lot more leadership involvement in that case. Again, we didn't apply the PII that frequently to his account, so the moments that that happened were few and far
between. But the instances in which his account was involved was a lengthier process
and involved a lot more senior leadership from the company than if it was another VIT,
someone who just had 100,000 followers and was an actor.

Q. Do you believe that made it more difficult for Twitter to take actions like
placing a warning on President Trump’s tweets during the election?

A. Yes, absolutely. And when you’re dealing with social media companies, and
the vector for harm is so small, things go viral in a matter of minutes, leaving something
up for multiple hours, in many cases, the damage is already done. One of the initiatives
that I worked on was an attempt to streamline this process for applying the PII to make it
clearer and to make it more expeditious, and it was shipped. And in some cases it
worked, but in other cases, as mentioned, if there was a tweet dealing with Donald
Trump, that play book, in many instances, went out the window and it would still take
quite a long time to reach a final decision.

Mr. Raskin. Jacob, could I follow up on that?

Please.

Mr. Raskin. I’m going to have to go in a moment.

We have other evidence consistent with what you’re saying, suggesting that
Twitter was bending over backwards at numerous points not to take down any of former
President Trump’s tweets, including tweets that would be taken down -- including tweets
that would have been taken down had someone else sent them out.

First I want to know, is that your sense as well?

The Witness. Yes, 100 percent, I think that Twitter did everything they can to not
take down Donald Trump’s tweets.

Mr. Raskin. And do you attribute that to the fact that he was President of the
United States? Do you attribute it to the fact that he had the most number of Twitter
followers, or do you attribute it to the sorts of things you were invoking before, lobbying or social connections? Where did that come from?

The Witness. I think -- and this is, obviously, my opinion. I think that it had to do with the fact that he was President of the United States and Twitter was terrified of the backlash they would get if they followed their own rules and applied them to Donald Trump.

Mr. Raskin. And this sense that you have and that others had, how did you arrive at that sense? How was it expressed to you that people should steer clear of taking down the President's tweets?

The Witness. I think by seeing numerous instances in which Donald Trump would tweet something and you would hear from senior leadership explanations of alternate interpretations that were, quite frankly, outlandish and not what the average person probably understood those tweets to mean. And seeing that on numerous occasions led, at least me to believe that, as mentioned, Twitter was bending over backwards to say, well, Donald Trump could have meant this and kind of inserting that interpretation of the tweet and using that as the rationale for why we did not take something down.

Mr. Raskin. In general your interpretive methodology was to determine how a reasonable person would understand a tweet?

The Witness. That's correct, yes.

Mr. Raskin. And so you weren't trying to torture out explanations for what the person subjectively might have meant in order to exonerate a particular tweet?

The Witness. Correct. It's quite difficult to try and understand someone's intent.

Mr. Raskin. Gotcha. Okay. Thank you very much.
I yield back.

Thank you very much, Mr. Raskin.

BY:

Q  Going beyond that, I want to ask if there were any instances where that happened leading up to January 6th where you particularly felt that a conversation with leadership reflected a desire to read in intent to President Trump's tweets that would not typically be given to another user?

A  Yes. And my memory is a little shaky. There are numerous instances. I remember one case -- and I'm not going to have all of the details exact, but it had to do with Donald Trump tweeting about the owner of "The Atlantic," and there was an article that was run that was, in Donald Trump's view, disparaging toward him and it had to do with veterans. And he tweeted something saying, like, let her know what you think, or something like that. And there was -- that was a case that I recall where it felt as though if it was not Donald Trump, we would have viewed that as potentially likely in violation of our policies.

And, again, I'm fuzzy on some of the details here, but that's an example, one of many, leading up to the U.S. 2020 election where it felt as though we were making different interpretations for the tweets that Donald Trump was posting.

Q  Thank you.

I just have a few more questions about the election more generally, and then I wanted to zoom in to coded incitement specifically.

So you mentioned that you felt that the safety policy team was not prepared to handle a lot of the threats that were sort of evaluated by the election threat model that Twitter had. And I wanted to ask if there were any particular projects that the safety policy team did undertake to evaluate election-related threats?
As I touched on earlier, there was the guidance that was intended to streamline and clarify the applications of our policies in a more consistent manner. There was the projects to try and streamline the application of the PII. This is all independent of coded incitement to violence, which we'll touch on later it sounds like.

But in general, it did not feel like there was any sort of organized analysis of our role leading up to the U.S. election. It was chaotic, and, again, leadership did not have a clear -- did not give our team clear instructions in regards to these are the things we should be looking out for, these are the things that we should be addressing. It was very much bottom-up work, and nothing I can recall beyond what I already mentioned.

Understood. That is helpful.

Before we move on, I want to ask about one meeting in particular that I believe you had on October 21st with other members of the safety policy team in Asia. Do you recall that meeting?

Yes, I do.

And can you give us an overview about what happened in that meeting?

Yes. So as mentioned, it was very unclear what safety policy's role was leading up to the election and how we should be preparing ourselves for the associated risks. And so, members on the team were asking leadership in that meeting, What is our North Star? What is our ultimate goal with the U.S. election and being prepared and responding to the risks? And it became quite contentious, and leadership became quite frustrated with the fact that folks were repeatedly asking this question. And to me, it underscores the fact that this was October 21st, this was mere weeks before the election, and colleagues of mine were still very unclear around what we should be doing, which is concerning.

So it was your sense that leadership actually was annoyed by staff's desire to
have a more comprehensive strategy towards the election?

A Yes. After the meeting, my supervisor called me and expressed frustration over the repeated questions of, you know, What is our North Star? So leadership -- that is a correct assessment that leadership was annoyed and frustrated that we continued to ask these questions.

Q And was it your sense that Twitter believed there were significant threats that were associated with the election? And by "Twitter," I mean Twitter leadership at this meeting.

A Yes. They were, I think, aware that there were risks, but it fell short of turning that into action or providing any sort of clear guidance. But I don't think it's correct to say they weren't aware that there were risks or they didn't think that there were risks. I think that they were aware, but they didn't do anything to help address those risks and concerns.

Q And in your mind what should they have done in response to that meeting?

A I think that should have been a signal that there was still a lot of confusion and questions around our role, and I think that it would have been helpful to have a team meeting and say, Here's what our staffing plan is going to look like leading up to the election. Here's how we're going to ensure that our communication is effective. Here's going to be our goal for addressing escalations as they come. You know, is that going to be a turnaround time of 10 minutes? 30 minutes? You know, I think leadership could say, Can you work with our operations team to find out the top five trends of what we're seeing in terms of problematic content?

I know this is all me giving ideas of ways that we could address it, but I think that some of those instructions and guidance and guardrails would have helped the team feel as though we at least have a sense of what our role is and how we can be prepared and
address those risks.

And to the point of communication, I think that was also a big aspect here and feeling like everyone on the team has the information they need to be successful in their role leading up to the election.

Thank you.

I wanted to say that counsel for the select committee, has now joined the Zoom.

But before we move on to talking about coded incitement specifically, I wanted to see if Representative Aguilar had any questions?

Mr. Aguilar. I don’t. I’m fine. Thank you.

Great. Thank you.

Q So at this point I wanted to pull up exhibit 1, if we could. And if you need a moment to review this, that is totally fine. We can scroll through it.

This is a draft of the coded incitement to violence policy that Twitter has provided to us from November 4, 2020, which is the day after the election.

So at first blush, do you recognize this document?

A Yes.

Q And so I -- oh, go ahead.

A Sorry. Yes, I do recognize this.

Q Great. And so, at the very top of the page, you saw that it had a date, November 4, 2020, and status: Shipped. Can you just explain to us what that means?

A Yes. So what that means is when it says "shipped," it means that leadership has approved this so we could start allegedly using this policy in the way that
it's intended, so actioning tweets for coded incitement to violence.

Q    Got it.
A    It's important to note here that just because something is shipped and it is
      approved doesn't always mean that that's -- that it is going to be used.  It basically just
      means that it's gotten sign off.  But as the case with coded incitement to violence, there
      were issues with it actually being used to take down tweets.

Q    And it's our understanding that between the Election Day and January 6th,
      there was not ever a clear consensus that coded incitement to violence could be used as a
      policy to action tweets.  Is that correct?
A    Correct.  If my memory serves, this was shipped, meaning that leadership
      had approved it, but there was a caveat that before we could actually bounce or delete
      tweets under coded incitement to violence, we would need those tweets to get one last
      approval.  And that's where we ran into a blocker in actually utilizing this policy.

Q    Understood.  And we'll get more into that specifically when that occurred
      shortly.  But first I wanted to ask you more generally, what was the coded incitement to
      violence policy, and what was it meant to accomplish?
A    Yes.  So, as mentioned, our team became aware of a gap in our policy.  So
      we have, as mentioned, 21 policies that we cover.  But we were aware of trends on the
      platform that we viewed as problematic and harmful, but we did not have a policy to
      address it.  And some of the tweets under this were "locked and loaded," "stand back
      and stand by."  And this policy aims to address content that is not clearly an incitement
      to violence, but as the name indicates, it is more coded language, but nonetheless,
      harmful.

Q    Understood.  So these were tweets that have potential to encourage
      real-world harm but did not clearly meet Twitter's standards of incitement or glorification
of violence?

A       That's correct, yes.

Q       And did Twitter have a policy on incitement to violence prior to this proposal?

A       There was much debate amongst the team around that. We had senior leadership attend a meeting saying, No, we clearly have an incitement to violence policy; but there were -- a lot of my colleagues expressed confusion as to what that actually looked like and what that meant. In a lot of the training materials that we have that you go through when you're onboarded, to my knowledge, the slide that was closest to incitement to violence did not go over this category specifically.

We had a policy around wishes of harm, which was close to this, and that was used later as a proxy, if you will, for inciting violence. But the fact that there was confusion around even a policy around inciting violence I believe underscores the lack of clarity on our team and confusion around the application of the policy.

Q       And can you give us a sense of when that meeting was that you had with leadership when they said there was an incitement policy?

A       I don't remember exactly. I believe it was between the election and January 6th, somewhere in that time frame.

Q       That's helpful.

So suffice it to say that you perceived a gap in Twitter's policies when it came to incitement, a lack of clarity, what could be actioned, and coded incitement was seen as a way to go back out?

A       That's correct, yes.

Q       And so, I noticed you mentioned phrases that would fall under this policy, "lock and loaded," "stand back and stand by." Those are, obviously, both phrases that
were attributed to President Trump in the final months of 2020.

So I wanted to ask you about the origins of the coded incitement policy.

A Exactly how we decided to write it, just --

Q Yeah. So, generally, like, what was the impetus for the safety policy team brainstorming this idea in the first place? Was there a particular precipitating event? I know you might not have been involved in all of those conversations, so if you were not involved, then that would be helpful to know as well.

Ms. Ronickher. And I'll just chime in here that also if there was any communication that would be potentially protected by Twitter's attorney-client privilege, you should not speak about that.

A Yeah, I think that the answer to that I can't speak to given what Alexis just said.

BY

Q Understood.

So I'm going to ask a couple more questions, and if the answer is the same, please just say so.

So is it fair to say that President Trump's rhetoric played a role in necessitating the need for a coded incitement policy?
[1:08 p.m.]

A  Yes. The policy, as mentioned, was aimed to fill this gap, and the gap that existed was by and large due to Donald Trump and his supporters, his base, and the rhetoric that they were using on Twitter. So, yes, it was in response to Trump and his supporters.

BY:

Q  If we can take the "locked and loaded" example. The coded incitement policy would have been used to help Twitter action responses to the President's tweets as well as the President's tweet itself. Is that correct?

A  Yes, that's correct. This policy could be used on any tweet on Twitter, be it from Donald Trump or a regular user, whether it's his supporter or not, it would be applied globally.

Q  And to generalize, it would be to lower the temperature of the discourse that was happening on the platform and minimize the potential for real-world offline harm?

A  The second half of that is true. I'm not sure if it's necessarily aimed at lowering the temperature. When we're discussing policies, I think our aim is to -- or was to at least mitigate any propensity for offline harm. So that is the aim here, where we were seeing this trend where it wasn't just one tweet, one person; we were seeing a movement starting and no way of addressing it. So that was exceedingly problematic. And coded incitement to violence was our way of trying to get at that and to try and reduce the potential for offline harm.

Q  And just to be clear, that movement you're referring to has a nexus to President Trump's tweets and what he was saying on the platform?
So I wanted to ask you about the other phrase you mentioned, "stand back and stand by," which President Trump said in his first debate in late September 2020, calling out the Proud Boys. Do you remember the impact of that tweet on -- sorry, that comment on your team of Twitter? Did it heighten your concerns?

It heightened my concern, and I think that's when we started to see this movement spin up. And as mentioned, it was not just one tweet, one person, one day when this was happening. We saw this gaining momentum.

As for leadership, Safety Policy did not have a staffing plan for when there were Presidential debates happening. There was not any sort of post mortem to understand as a result of the debate and something that might have been said, do we need to readdress our approach to the U.S. election. From leadership's end there was no communication that stand back and stand by made them more concerned about risks leading up to the U.S. election.

But from your vantage point, you did see it change, I believe you just said, in how users were posting following the President's comments at the debate?

Yes.

Can you answer, then, what you saw on the platform?

Yeah. I think that this was kind of the beginning of the movement that ultimately led to January 6th. So you get Trump saying this, and it emboldens his base. And we saw that on certain tweets. And that, again, kind of was the reason why we had this coded incitement development policy, where we were seeing more and more of this content and no way to address it.

So at least from my vantage point, there was concern when you're getting escalations, be it from other cross-functional stakeholders, other teams, and there's no
way to take it down under our existing policies. So it was the beginning of what gained
momentum and likely led to the attack on the Capitol.

Q And so if we can draw down on that a little bit. Users taking the President's
comment, sort of running with it on a platform, what did that look like? Was there
cause to actually stand by, to mimic his language exactly, or were there other themes that
had been bubbling up about sort of potential for locked and loaded? But did these
themes coalesce or was there a more common response to this tweet above?

A That's a good question. So, as you mentioned, our team is not evaluating,
you know, a large swap of tweets on a daily basis, so I can't necessarily speak to exactly
what was happening. But there were, I would say, increased -- there was a spectrum
where there was severe tweets that were clearly in violation of our policies if there was a
specific violent threat. But I would say, in general, this seems to be a turning point of
the President in many ways feeling like there was this tenor of violence, and that, I think,
is what was particularly concerning. And so you would see more of these come up on,
be it escalation, or we would be able to use some of these examples in the development
for coded incitement to violence policy.

Q Got it. So it was certainly a case in point for why this coded incitement
policy was necessary in the view of your team?

A Yes, that's correct.

Q Got it. That's very helpful.

So in the month between that comment and the draft of the policy that we see
here, a month and a few days, that seems to be, based on our conversation so far, a
pretty major shift in how Twitter was perhaps willing to treat President Trump's account.
Would you agree with that?

A Yeah. And to be clear, the policy, as mentioned, was to apply to anyone at
Twitter.

Q  Uh-huh.

A  So it wasn’t specific to Donald Trump, but it was specific to the issue at hand.

But, yes, I do agree this was a shift and hopefully aimed to address some of the concerns and risks that we were seeing.

Q  So what was Twitter's general posture towards the President's account prior to this point, October, November of 2020?

A  When you say Twitter, do you mean senior leadership?

Q  Yes, yes.

A  Yes. I mean, to be clear, I think that senior leadership, even after coded incitement to violence, was reluctant to toss Donald Trump's account. We were seeing -- I was seeing more and more escalations that involved Donald Trump's account.

As mentioned, the first time we applied the PII was in the summer of 2020. And after that point, we were seeing escalations from Trump on a pretty regular basis, and leadership did not seem to want to take action, even though we saw the tweets and the tenor was getting more hostile, there was this undertone of violence. But there was, again, a reluctance to do anything to his account specifically.

Q  And so would you say -- you mentioned that although this policy, this coded incitement policy was finalized, you often ran into problems getting final approval to remove tweets under the policy. Is that related to the reluctance to take action against accounts that were related to the President's account in some way?

A  I think that -- I mean, it would be conjecture at this point to try and understand why there was not signoff from leadership to remove these tweets. We had compiled a spreadsheet with hundreds, if not thousands, of examples of tweets that fall under coded incitement to violence, and we never got final signoff to remove these
tweets. And I think, in my opinion, it goes back to the fear that senior leadership has of the reputation that they are stifling voice on the way and being -- you know, the potential that they get sued, the potential that they lose money.

I think that, you know, they can say that it's about promoting voice and promoting speech, but when we see the speech is clearly problematic, and that's -- you know, I think can be pretty massively agreed upon. Yeah, I think that there was a reluctance there to be seen as biased. And so they used the angle of safeguarding speech as the reason for not doing this, but I don't necessarily totally believe that.

Q And these tweets that you collected that were, in your view, in violation of the coded incitement policy, they were mimicking the President's language with phrases like "stand back and stand by" and "locked and loaded"?

A Yes, yes. They were all mimicking the President's language. And there was, as mentioned, an intense spreadsheet, and that spreadsheet was not comprehensive of all of the tweets that existed out there. That was only from a certain point in time. And we had leadership actually go in that spreadsheet and look for pretty much any reason other than coded incitement to violence to potentially remove a user. But we never got signoff to -- "bulk action" is the phrase -- bulk action any of these tweets.

Q Understood. And we'll get into a little more of that timeline soon.

I did want to ask if you wanted to take a break. We've been going for a little over an hour. But first I wanted to see if anyone else had any questions on this portion. or ?

I see none. Let's take a break until 1:35, if that works for you all, Eastern.

Ms. Ronickher. That's fine.

Great. Okay. Thank you so much. We'll go off the record at 1:21
p.m.

[Recess.]

So, with that, we can go back on the record at 1:35 p.m.

BY [Redacted]:

Q And so at this point I wanted to continue our conversation about coded incitement by bringing up exhibit 3, which you could take a minute to look at. This is a timeline that Twitter prepared for the select committee as a way to summarize the steps that it took throughout the election season and leading up to the attack on the Capitol on January 6th.

So if we can scroll to the bottom of page 14. There are two entries that I want you to look at related to the implementation -- or alleged implementation of the coded incitement policy. So it's page -- no. [Redacted], we can scroll back up because it's page 14 in the document. No. Yes, perfect.

So, as you can see there, that -- Twitter writes that the safety policy team develops guidance that could be actioned under coded incitement. And then under No. 4, the safety policy team updates coded incitement guidance.

So is it fair to say that the coded incitement guidance was indeed developed and updated but never actually implemented?

A I'm sorry, I'm just continuing to read this.

Q Sure, sure. Go ahead. If you want us to scroll, let us know.

A Yes. So this is technically accurate, but it omits material information that the tweets that would fall under this guidance were never actioned.

Q So when Twitter writes that on Election Day and after Election Day the safety policy team was working on updating and developing a policy that would allow these inciteful tweets to be taken down, it is the same policy that Twitter leadership refused to
let your team implement?

A That’s correct. Yes, this is the same policy that we collected on the spreadsheet hundreds of thousands of tweets, and we were never given the approval from leadership to actually do anything with those tweets. We were never allowed to remove them from the platform.

Q And this was despite reports of coded or opaque cause to violence on Twitter increasing?

A That’s correct, yes.

Q And the safety policy team asking for permission to implement this policy?

A Yes, that’s correct.

Q So you just mentioned your effort to collect some of these tweets. And actually we can go to the top of page 15 where there is an entry. So here you go. It’s the same entry, it says from November 4th to November 9th, Twitter works to surface and review tweets that might include current language or the potential to incite violence.

To your mind, does this represent the effort that you and your team undertook?

A Yes, this does represent the efforts of the team. And, again, this just does not include detail that I believe is crucial that says these tweets that were surfaced and reviewed were never removed from the platform.

Q So Twitter including this as part of its summary of actions taken around January 6th, while technically true, omits the fact that none of these coded tweets were allowed to be removed by Twitter leadership?

A That’s correct, yes. Throughout this time, from the election to January 6th, in nearly every team meeting or in our team staff channels, we had repeatedly asked leadership for guidance on how we should treat this type of content. In addition to the
spreadsheet where we had corrected and surfaced tweets that we believed violated this policy, we were also receiving numerous escalations, often on a daily basis, about other tweets that were reviewed under coded incitement to violence, and we never received guidance from leadership around how to action this type of content.

Q Thank you.

So the time period we're talking about here, November 4th through November 9th, that was the initial effort immediately following the election to correct, you said, it was hundreds of tweets in a spreadsheet, right?

A Yeah.

Q And was that an unusual effort for the safety policy to undertake to kind of present leadership with that many tweets in that way?

A Yes, yes. During my tenure at the company, I had not seen a similar instance where we coded the new policy, surfaced and reviewed, you know, thousands of tweets and then asked for leadership approval. So, yeah, this was a unique instance.

Q So the thousands of tweets that you reviewed that Twitter talks about in this entry, that was -- the review occurred because there was no clear guidance on how to use this policy?

A That is correct. Well, let me just clarify. So we created this policy, and we worked with our operations team to surface and get at this content that met the criteria in the policy language. So we had a spreadsheet that we were able to proactively review and get content. And then there was addition to that, the reactive side, where we were not proactively looking for this content, but it was escalated to us as problematic. And in both instances we were not given guidance on what to do with this.

So the spreadsheet, we never got approval to take it down. For instances where we were reactively getting tweets, be it from other stakeholders within the company,
external partners, there were various avenues in which we would get a tweet escalated to us. There was maybe one or two instances in which leadership decided to remove one singular tweet. But, again, it goes back to the issue of something being or decisions being arbitrary and our team having no clear understanding of how to address this type of content.

We had a clear answer on the spreadsheet, but we were not allowed, we were not given the approval to bulk action or bulk delete these tweets. So we at least had an answer there that leadership said, no, you cannot remove these. But for the reactive tweets that were on escalation, we still had no understanding, was this just a straight out, no, we’re not removing this; should we sort of validate this as something that we should remove from the platform? It was extremely unclear with how we were dealing with this content.

Q That’s very helpful. So before we get into the actual meeting that I believe resulted in the order that you could not bulk delete these tweets, I wanted to ask you about -- thank you, [redacted] -- I wanted to ask you about some of the content that you were seeing as you were undertaking this review of thousands of tweets.

And this is immediately after the election. President Trump refuses to concede once the election is called in that November 4th to November 9th timeframe. What were you seeing that was concerning to you?

A Well, we were seeing Donald Trump saying that there was election fraud. He was tweeting very regularly. And as a result of that, we were seeing users on Twitter echoing his sentiment that there was election fraud. And then in addition to that, we started to see this movement gain increasing momentum for stand back and stand by, "stop the steal." And so this content was becoming increasingly more common. We were seeing it pretty much -- every escalation that we were dealing with had to do with
this type of content. And, you know, we were just starting to see this undertone of rhetoric before the election. It was becoming stronger and stronger leading up to January 6th.

Q You believe that this kind of rhetoric further gained momentum when President Trump refused to concede on November 7th?

A Absolutely.

Q And you said it was your work -- job, the job of your team to prevent offline harm. So it was your belief that this rhetoric was seen as a potential to create violence on the ground?

A Yes, that’s correct. And I think what made this a unique situation is the magnitude that we were seeing this content proliferate. So as I mentioned earlier, it wasn’t one tweet or even a series of tweets; it was thousands of tweets we were seeing. And we were not allowed to remove them from the platform. And I think it was because there were so many tweets, there was indication that things would become violent.

Q Can you scroll down a little bit more?

There were so many tweets that had this undertone of violent rhetoric that your team believed it was likely that there would be some kind of offline activity?

A Yes. What I’m trying to say is I think that there is a difference between one account saying "stop the steal," "the time to stand back and stand by is over," you know -- you know, a series of accounts doing that. What I’m saying is where I and my colleagues were very concerned about this was that it wasn’t just one account or a few accounts, we were seeing this entire movement happening and no way to address it. And not only were we seeing this movement, we were seeing it gaining momentum over the course of many weeks and months.
Q And it was your impression that President Trump's activity on Twitter during
that time helped give that movement momentum?
A Yes, yes. Absolutely.
Q Go on.
A There was another team that dealt with our election integrity policy and that
had to put labels on Trump's tweets that claimed election fraud. And there was a point
in time where I remember my inbox had hundreds of emails saying we have labeled
Donald Trump's tweets. And it was just numerous instances of labeling Trump's tweets
about election fraud. And, to me, this shows that he is putting his message out on
Twitter and he's doing it continuously.
And so, yes, it was integral to what his users decided to ultimately do and say on
Twitter and the actions that they took on January 6th.
Q Thank you.
So I want to ask now about a meeting that I believe the safety policy team had
with the vice president of Twitter, Vice President Del Harvey, to present her with the
collection of tweets that it believed violated the coded incitement policy. Were
you -- are you aware of that meeting?
A Yes, I am.
Q Did you attend that meeting?
A I believe so, yes.
Q Would it make sense to you that this meeting occurred on November 9th, per the Twitter timeline?
A Yes, that makes sense.
Q And do you remember the result of that meeting or what happened during
it?
Yes. It's coming back now. This was a meeting where Del was providing different interpretations for the content that we deemed was coded incitement to violence. I remember that she was advocating for another interpretation for "locked and loaded" and how that could be someone in their house acting in self-defense. And this was a meeting where we ultimately, again, chose to not remove the tweets that we collected in that spreadsheet.

So as I mentioned earlier how there were interpretations of content that we felt were not realistic or reasonable to the average user, where we did not think that "locked and loaded" meant someone was acting in self-defense, particularly when you could see that there were from other profile elements -- like Trump supporters clearly angry about the results of the election -- it was unrealistic to believe that it meant something that wasn't going to be violent and them initiating it. And being self-defense seemed like a silly explanation of what that tweet meant.

So you didn't believe that Del Harvey's interpretation of the tweets were particularly credible, it seems?

Yeah, that's correct.

And I think that day she was the one who told you that your team was not allowed to action any of these tweets?

That's correct, yes.

And why do you believe that happened, from your vantage point?

I think it goes back to this fear that Twitter has of being perceived as bias, being perceived as a left-leaning liberal company that is stifling right-leaning speech. I don't think that she likely viewed the risks as being as real and as severe as they ended up being. At the end of the day, I don't know why she chose to not remove those tweets.

I think that this policy was different in that it was coded, right? So it was not as
categorical, as explicit, and that could be an argument as to why content wasn't being
removed. But as I mentioned, when you take the context into consideration and review
the totality of what was happening on the platform, it was hard to ignore the propensity
for offline harm.

Q And so if I'm understanding what you're saying correctly, it seems like
because a lot of these tweets were coming from President Trump supporters and in
response to statements and tweets he was making after the election, there may have
been reluctance to take action against that coded incitement because it was coming from
a particular segment of the American political discourse?

A Yes.

Q And that was in response to some of the external pressures Twitter felt
about its own reputation for bias or unfairness?

A Yes.

Q Understood. So at the end of the day, you didn't have the authority to
remove those tweets you collected after November 9th, that November 9th meeting,
correct?

A That's correct, yes.

Q And after that point, did the safety policy team continue to ask for
permission to implement the coded incitement policy in the weeks after the election?

A Yes. Nearly every day we asked for the ability to remove this content. We
asked for clear guidance on what we should do with this problematic content, and never
received an answer.

Q And this was because you continued to see a worsening trend of violent
rhetoric or rhetoric endorsing violence on Twitter?

A Yes. As mentioned, there was that spreadsheet where we proactively
surfaced problematic content under coded incitement to violence. But in the weeks following, we were also receiving a barrage of escalations that dealt with this type of content. And, again, we had no approval to take it down. We had no clear guidance on how we should be treating that type of content.

Q And so did you become more worried that there would be offline harm in, let's say, by late November when you had seen now weeks of this rhetoric on Twitter?

A Yes, yes. And when Biden was declared the winner, we saw that there were Trump protests and, you know, the intensity and heat did not turn down when that happened. And so there was a concern on the team of what is the tipping point going to be, because it did not seem as though it would go away on its own, particularly when Trump is continuing to tweet that there was election fraud. He is surfing the flames of this fire here.

Q So your team, and you particularly, saw a desire on Twitter to violently resist -- or endorse violence to resist the outcome of the election once Biden had been declared the victor?

A Yes. We were continuing to see more and more content around things turning violent. Again, as mentioned, an increase in escalations pertaining to, you know, the time to stand back and stand by is over, and the rhetoric was increasing in intensity.

Q And as this -- just to be clear, as this rhetoric of civil conflict, violent disruption of the election increased, you weren't able to remove any of those posts?

A That's correct. We were not able to.

Q Understood. So given that, my last question on this particularly is about what Twitter told the select committee in that document, given how your authorization was curtailed to act under the coded incitement policy, what do you make of what Twitter declared to the select committee that these tweets were surfaced and reviewed
and that there was an updated guidance on coded incitement and failing to tell the other
half of the story that you've just told us?

A I think it's an attempt for Twitter to make themselves appear like they were
responding to the content and risks on the ground. And I think that it's just categorically
false, given what I saw. As mentioned, it's a clever way of going about it where it is
technically true that there is material information that they are not including in that
which changes the narrative.

Q Thank you. So I have other questions on the timeline, but before I do that,
I wanted to see if anyone else has questions on this piece of it.

But seeing none, if we can pull back up exhibit 3, kind of just run through a few
more questions on it. So, [blank] if you can scroll to page 4. Page 4 of the
document. It's confusing.

So here you could see -- actually, if you can scroll down a little bit more. Perhaps
a little more. Here we go.

Towards the bottom, you see this is Twitter talking about some instances where it
applied a label to President Trump's tweets. The first is related to election
disinformation, and the second on the next page is related to glorification of violence?

A Uh-huh.

Q And that's related to some of the protests after the murder of George Floyd.

So we talked about these earlier, but I wonder if you have a reaction to Twitter
including this on the timeline whereas the election -- given some of the reluctance we've
talked about on Twitter's part to action the President's tweets as we got closer to the
election and then afterward.

A Yes, I think that Twitter is trying to make themselves seem like they were
responsive to the risks; that they were being proactive; like, they weren't afraid to take
action on Donald Trump's account. The timeline does not include the numerous
escalations where we did not label Trump's tweets, where they were -- could easily have
been considered in violation of our policies.

So I think, to me, it is -- they are choosing to include specific elements in their
timeline to make it appear as though they were being potentially more proactive and
taking more action than they actually were. I think that it does not paint the full picture
of what was actually happening during this time.

Q Thank you. That's helpful.

So I'll ask you one more question. This also relates to what we talked about
earlier with the interstitial warning messages. And it's on page 10 and --

I'm sorry, I think it's page 10 of the actual document.

Okay. So here we go. On October 9th, there's the launch of a number of
election-related initiatives. And I can give you a second to read that if you want to.

If you could scroll down a little bit, .

So what I was interested in is the expansion of the use of the interstitial tool to
apply tweets from U.S. political figures. And it would prevent likes, retweets, or replies
to certain tweets by political figures. And I wanted to ask if that was the same
interstitial initiative we were talking about with President Trump's tweets earlier in the
year?

A Sure. I'm going to take a minute to read this.

Q Yes. Please do.

A Yeah. So they are claiming that this was an expansion. This was on
October 9th they said that this happened?

Q Yes.

A As it relates specifically to the interstitial, this to me is what the interstitial
did prior to October 9th. I don't see this as any sort of additional initiative to mitigate offline harm. The interstitial was, again, a click through morning over a tweet. That was not anything new. That was what was used in May of 2020, June of 2020. And, yes, this was something we would not be able to retweet it as well.

Q But as you were saying, certainly in the case of President Trump, when something was blocked with the interstitial so it couldn't be retweeted, you saw users mimicking that language in their own tweets over and over again. So it wouldn't have necessarily had the impact that's described here, it seems?

A That's correct. And you can see they say here, in addition, users are only able to quote tweet. And so that was what we saw where it was almost amplified in the sense that if we were to apply the PII to one of Trump's tweets, sure, you might not be able to retweet it, but his 80-plus million followers, many of them would quote tweet exactly what he said. And we were not equipped to proactively detect all of those quote tweets and remove them from the platform. We can try to get at some of them, but, again, there was no process in place to address all of the quote tweets.

Q And just to clarify for the record, quote tweeting is essentially retweeting with a comment of your own?

A Quote tweeting would be essentially copying what that person said and tweeting it.

Q Okay. I guess my understanding of Twitter is not as good as I thought it was. But all right. Thank you. That's helpful.

So because you couldn't necessarily limit the spread of the content that was being posted in these tweets, this sort of step here, this initiative was not, in your mind, sufficient to tackle what you were seeing on the platform with President Trump's content leading up to the election?
Correct. I do not think it was sufficient.

Okay. Thank you.

So we can leave it down now. Actually, if we can go to exhibit 2 instead.

So this is another document provided to us that is the post-election protests and calls for intervention guidance, November 5th. Are you familiar with this document?

Yes.

Okay. So can you explain the significance to us briefly?

Yes. I believe that this document was [inaudible] in to provide guidance to our operations team to address content that, again, might lead to offline harm in the environment that we were seeing post-election. A lot of this was -- if memory serves, leadership created this almost behind closed doors where there was not a lot of input or clarity from the safety policy team specifically.

So that would actually bring me to my next set of questions which is about whether or not the coded incitement policy was included in this, because you had a document we just looked at which was from the day prior where you had a coded incitement policy that had been finalized and approved by leadership at least technically, correct?

Yes, yes.

So this document says "approved" at the top rather than shipped. Can you just briefly tell us the difference between those two statuses?

There was not a unified set of terminology for a document. I think that some people would put approved, some people would put shipped. I don't know if there's a ton of significance on the two.

Okay. But at the end of the day, there would be sort of additional steps taken to allow staff to act on a document like this?
A: That's correct, yes.

Q: And this was the universal rules that the operations team and others were using to actually deal with tweets in the post-election period?

A: Yes, that's correct.

Q: Okay. So if we can go to the bottom of page 8. I hope that it's here. Great. So here at the bottom of page 8, you can see some language about incitement with election processes. And I'm particularly interested in those last bullet which seems to summarize sort of wish of harm policy. Is this reflective of the coded incitement policy or a narrower set of content?

A: Would you mind repeating that last bit?

Q: Sure, sure. So I'm curious about your view of what this last bullet accomplishes. Does it incorporate the coded incitement policy or does it simply restate other Twitter policies like the wish of harm policy?

A: Oh, this does not incorporate coded incitement to violence. This is our standard language that was actually taken from our Black Lives Matter playbook and was repurposed for this post-election period. So as you'll see, a lot of the examples here, the themes are more for Black Lives Matter than necessarily post-election period.

Q: So is it your belief that coded incitement occurs when a policy was incorporated into this post-election guidance?

A: No. According -- specifically this last bullet, that does not contain coded incitement to violence. That is our standard language for our wish of harm policy.

Ms. Ronickher: Mr. Glick, I think it might be helpful if my client could kind of review the document. I know it's a little bit hard on this, but to see --

Sure. We can go off the record --

Ms. Ronickher: Sure.
We can go off the record for 5 minutes. Off the record at 2:11.

[Discussion off the record.]

Ms. Ronickher. All right. We appreciate it. Back on the record.

Great. So we can go back on the record at 2:17.

Q So I'll ask the witness, having reviewed the document again, is it your impression that the coded incitement to violence policy was incorporated into this post-election guidance?

A Yes, it looks like there are elements of coded incitement to violence. In this -- in my opinion, it doesn't clearly get out the content that we were seeing, but the words are in there, yes.

Q But as we discussed, there was still no clear mandate that Twitter employees were able to action tweets because they violated the coded incitement?

A That's correct, yes.

Q And would it have been better, in your view, if the coded incitement had been approved and operationalized as a policy on its own instead of sort of sprinkled throughout this document?

A Yes, yes. With having been approved, there were still numerous questions and confusion from the team about how to treat this type of content. So if it was a standalone policy with clear guidance on how to treat content that fall within it, we would have had a much better way of getting at the problematic content we were seeing during that time period after the election.

Q Understood. So we don’t have to pull it back up, but I will read out a sentence from the Twitter timeline that we just reviewed. And when it's discussing -- thank you,
When it's discussing this post-election guidance, in an entry, Twitter writes to us that this guidance, quote/unquote, integrates and supersedes the coded incitement guidance from November 3rd. So if it supersedes the coded incitement policy, does that essentially mean that there was no clear coded incitement policy because it was not clear -- clearly stated in the text you just reviewed?

A Yes, that's correct. And as I mentioned earlier, that document was created in a silo with leadership. And it was not clear to myself, my colleagues, what the approach was with the content that we had been seeing. And as mentioned, there was confusion if we had an incitement to violence policy. There was confusion particularly around coded incitement to violence.

That document -- again, it has the words in there, but that doesn't necessarily mean that it has clear guidance and understanding from employees implementing that on how to actually apply that, get at that content that we were seeing.

Q So was it your impression that the policies that you just reviewed clearly stopped the gaps in Twitter's incitement policies or did it restate the policies that already existed which already had significant gaps that we discussed?

A It restated the policies we already had and did not address the acts that we were seeing.

Q So were you concerned when you reviewed this code -- this post-election guidance and saw that it did not include clear guidance on coded incitement or any other clarity on incitement policies?

A Yes. And for what it's worth, we were not given the opportunity to review or provide input on the document before it was approved.

Q And you said that some of the language was lifted from a prior document related to protests over the summer. Is that correct?
Yes, that's correct.

Did you believe that it was appropriate to replicate the model of sort of Twitter's policy posture or should there have been something else going on?

I do not think that it was appropriate because it did not adequately address the risks that we were seeing and the content that we were seeing. The situation where there were people protesting the death of George Floyd was drastically different than what we were seeing during the post-election period. And the policies that we were using during that time, during Black Lives Matter were more -- there were less gaps, I would say, than when we were seeing the certain rhetoric that was being used during the post-election period. There's less of this coded language during the summer of 2020 than we were seeing in the post-election period. So, no, I don't think it was appropriate to copy and paste from that guidance.
And to drill down on that point a little bit more, one of the differences, it seems like what you're saying, is that there was consistent involvement by President Trump's own tweets in this sort of coded rhetoric version on violence or civil disorder that was the centerpiece of what you were seeing after the election, correct?

A Yes, that's correct.

Q And your response to that, just to repeat it again, was the coded incitement policy?

A Yes, that's correct.

Q And that was ready to go as of Election Day?

A Yes.

Q But Twitter leadership instead chose to put out this document that you did not have the chance to review and that Frankensteined together pieces from earlier crises?

A Yes, that's correct.

Q And in doing so, it sidelined -- Twitter leadership sidelined the coded incitement policy that your team, the team of experts, had created?

A Yes, that's correct. The guidance, the post-election guidance as mentioned, it has the words "coded incitement to violence," but the substance of the guidance itself does not adequately address the problematic content that we were seeing on platform during that time.

Q Thank you. That's really helpful.

So at this point, I want to ask a couple of questions about the run-up to
January 6th, and hopefully we'll get kind of up to the day of January 6th, and then we can take another break and then have one more session, if that works for you.

A Yeah.

Q Great.

But first, does anyone have any questions from the staff?

Seeing none, we can move onward towards the more immediate pre-January 6th period once we can clear that you weren't going to have a coded incitement policy at your disposal.

So I wanted to ask, sort of generally -- and we touched on this before -- but in your view, what was the implication and the consequence of not having this coded incitement policy as you moved from the November 9th time frame towards Thanksgiving and into December? What was happening on the platform?

A I think the movement was becoming stronger. It was becoming bigger. We were seeing the situation intensify. We were seeing threats against election poll workers. We were not seeing the situation resolve itself, or seemingly slow down in any way.

Q So what was most concerning to you? You mentioned election workers. I am curious if you saw discussion at this point about plans to come to Washington at some point to protest the election results?

A In December, Trump's tweet about "Be there, it will be wild" was obviously concerning. And I think for me, the thing that was most concerning was there was no appetite from leadership to address this problem. It was as if they had their hands tied and they couldn't do anything, and it felt as though we were helpless in trying to do something, anything to get this content off the platform.

Q You mentioned the President's December 19th "Be there, will be wild"
tweet. What was the importance of that tweet, in your view?

A It gave a time. It gave a place. It gave an opportunity for this movement to come to a head. On our team I think myself and my colleagues -- I'll just speak for myself -- we saw it as a very ripe opportunity for violence to break out, given what we had been seeing on the platform for months at that point.

Q So prior to the December 19th tweet, had you been reviewing a lot of content about the possibility of a protest on January 6th at the Capitol specifically?

A I don't remember specifically whether ahead of that tweet we were reviewing content specific to January 6th.

Q But what about after the tweet?

A After the tweet, yes, we were. I think that that tweet in many ways kind of crystallized the plans, and it gave his supporters a place to do what they were going to do. And it gave for us, you know, a time where we should be prepared and try and be on the lookout for anything we could do from Twitter's perspective to try and mitigate the potential for harm there.

So, yes, we were aware of that. I think that, you know, there was -- in the weeks leading up to January 6th, there were multiple dates that were of concern for us, be it the inauguration. Obviously, after the tweet in December, January 6th popped up. So those were all dates where we were concerned that violence would break out. And, again, leadership didn't seem to have any major concerns.

Q Got it.

So you mentioned that after the December 19th tweet, you did start to see Twitter being used to talk about movements to D.C. Is that correct?

A Yes.

Q Was there ever any discussion of whether or not users should bring firearms
to D.C. or convey D.C.'s gun laws?

A I don’t remember specifically about that. What I can say is, again, it became more specific after that time, where prior we were seeing folks, users on the platform often speaking in more general terms. And then after that it did become a bit more specific and honed in on January 6th and D.C. But specific to bringing firearms and evading the law, I can’t quite remember.

Q That’s fine. That’s very helpful.

So did you have discussions with Twitter leadership or your supervisor about the December 19th tweet in particular?

A I don’t entirely remember. I think that I remember we spoke amongst ourselves about the tweet, but I don’t remember if there was senior leadership involvement in that tweet. But it was a point of conversation given the concern of that date specifically.

Q And in your internal policy conversations on the safety policy team, was there a sense of this tweet drove home the need for a coded incitement policy?

A Yes. It underscored the need for it, and it furthered our team’s frustration about hearing nothing from leadership in regards to how to respond to this and still having no clear guidance on what to do with the multitude of tweets we were seeing that we knew had the potential for offline harm.

Q So just to draw an example, there were responses to the President’s tweet which identifies the time and place, as you said, that uses a phrase, such as "locked and loaded," or "standing back and standing by," you wouldn’t have been able to remove those tweets? They could still be up on the platform --

A That’s right.

Q -- until the President’s suspension?
That's correct, yes.

And you would have been able to if the leadership had let you act on the coded incitement policy that had already been approved?

Can you say that one more time?

You would have been able to remove some of those tweets if leadership had allowed you to act on the coded incitement policy?

Yes, yes.

So I wanted to ask more generally, what was your supervisor and Twitter leadership's posture towards implicit incitement and monitoring threats during this period when you brought concerns to them?

They were very laissez faire. I don't think that they took it as seriously as we were taking it, as I was taking it. Again, as mentioned, there was that meeting back in late October where my supervisor got frustrated about us continuing to ask questions about the North Star, and that trend continued when we were highlighting the risks post-election where it felt as though leadership was fed up with our team continuing to ask what we were going to do, ask what our plan was, seeking guidance on this content. It felt as though leadership thought that we were being overly dramatic with raising these concerns, and, again, never -- never seemed to take it seriously.

And we talked earlier about your belief that the safety policy team was not adequately resourced or particularly well managed leading up to the election. How would you describe the management of the safety policy team during the weeks between the election and January 6th?

The same. I think it was very poorly managed. I think that given the fact that it was over the holidays, there were a lot of people who were taking time off, who were not monitoring the platform. At one time, I was the only person who was on
safety policy online covering for the entire globe. So there was not any major concerns from leadership during that time.

Q So just to go back to that point, you said you were the only employee covering safety policy, which is the most controversial -- the review of the most controversial borderline tweets that have the potential to create offline harm, correct?

A Yes.

Q So can you give me a sense of how many tweets you might have been monitoring in the U.S. on any given day or expected to respond to?

A So I would be expected to respond to anything that came to our team. And it's variable, depending upon the day, the time, how much comes to our team. And I think during that time, as mentioned, there were just a lot of folks not working over the holidays, not monitoring the platform, not sending content to review; but I would be responsible for responding to anything, be it from the United States or elsewhere in the world to evaluate.

Q Understood.

So I want to circle back to ask about whether or not you believed your supervisor was attentive in meetings about these kind of threats or engaged in the need to respond forcefully. It seems like the answer is no.

A Correct, the answer is no. I think there were oftentimes where I would note my supervisor would be not paying attention, on their phone. There have been times where I've noticed sighs or eye rolls when we would continue to bring up the topic of risks around the election, coded incitement to violence.

In team channels we would say we're, once again, asking for clarity around how we should treat this. We just got another escalation. What should we be doing? And it would often be no response. We might get a response a couple of days later that said,
Oh, I spoke to Del about this. We're leaving it up without any further information as to why or what that means for other tweets we're seeing, if we should take that logic and apply it to other content. There was a total lack of transparency and information.

Q  And this was primarily sort of asked by your team about implementing the coded incitement policy, correct?

A  Yes, that's correct.

Q  Did you feel like your supervisor was adequately trained on some of the other key policies that the safety policy team had in its toolbox?

A  It depends on what you mean by "adequately trained." I believe --

Q  Maybe "familiar" is a better word.

A  Oh, familiar, yes, yes. You know, went through training and received training, so was familiar with the policies but, in my opinion, was not as well-equipped to understand the nuances or apply them.

Q  And did you ever request any action be taken to remedy this dynamic?

A  Can you say that one more time?

Q  Sure, sure. Sorry.

Did you ever request action be taken to remedy this sort of dynamic?

A  Yes. I spoke to employee relations numerous times, starting in the fall of 2020 through early 2021 -- this was even post-insurrection -- about some of the deficiencies I saw on leadership on safety policy. I compiled a document with numerous pieces of evidence that details my experience with some of this lack of leadership, lack of understanding of our policies. And employee relations told me that my supervisor might go through retraining. I don't know if that ever happened. In my opinion, the situation never got better.

And also, you know, post the insurrection, I told employee relations that because
of what happened, employees are going to start leaving the team, and I was not taken
seriously. And then multiple people left, including myself.

Q So when you were going to employee relations for the attack on the Capitol,
part of your concern was driven by this failure to engage in some of the threats of
potential offline harm related to the election, the certification of the election, that your
team was concerned about?

A Yes, yes. I think just totally not understanding the policies, the implications
of them. A big component of this was timely communication where, as I mentioned, the
vector for harm on social media companies is really small. Things go viral very quickly.
So I might assess a tweet as violative, but if I don’t get approval from my supervisor in a
timely manner, my work doesn’t necessarily have as much impact if that tweet is left up
for multiple days.

Q Got it. And so one more question on this.

Between December 19th and January 5th, the President tweeted over a dozen
times about the event on the 6th. I am curious about the further acceleration of some
of the concerning trends we’ve talked about during that time frame that you saw, and
also, whether or not leadership had a response to concerns that you were raising in late
December.

A Leadership did not have a response beyond keep doing what you’re doing,
which was -- you know, we were trying our best. But, no, we did not have a response
from leadership during that time really at all.

Q But in your view, the content kept becoming more indicative of potential
violence?

A Yes.

------ Thank you.
At this point I was going to talk next about the January 5th meeting, the night before the attack, but I think it might be a good time to take a break, if that works for you all.

The Witness. Yeah.

Great. It's 2:42. If you want to come back at 3:00 and take a little bit of a longer break, we can just do that. That sounds good?

The Witness. Sure.

Great. All right. Thanks so much.

Ms. Ronickher. All right. Thank you.

We can go off the record at 2:43.

[Recess.]

With that, we can go back on the record at 3:00 p.m. Eastern.

Q I want to move now to a discussion of a meeting that the safety policy team had on January 5, 2021. So the witness, if you could tell us, were you present at that meeting on January 5th?

A I was, yes.

Q And who attended that meeting?

A It was safety policy members from Europe, our Dublin office specifically, and our U.S. team members and leadership on that team. So we had a supervisor in Dublin as well as a recent hire supervisor in the United States.

Q So the supervisor in Dublin was in charge of the meeting?

A Yes, that's correct.

Q But your typical safety team -- safety policy team supervisor was not present on the call?
That is correct. There was, I believe, a death in the family.

So was this a regularly scheduled meeting, or had it been convened specifically to deal with what was going to happen the next day?

This was a regularly scheduled team meeting.

And could you give us an overview of what happened during that meeting as it relates to January 6th?

Yes. A colleague of mine asked our supervisor in Dublin, yet again, what we should do to prepare and plan for the rally at that point we knew was happening on January 6th. My colleague expressed concern over the trends we have seen on the platform over the previous weeks, months and, yet again, asked for guidance and clarity around how we should be treating such content and expressed concern around the highlight we had for violence that might happen the next day.

So at this point, there had been a further escalation of some of the content that was specifically suggestive of violence on January 6th?

Yes. As mentioned, we had continued to see momentum increase and more content being escalated to us and no guidance in regards to how to treat that content.

So when your colleague made this request for additional clarification on how to deal with this content, their chief ask was essentially to have the coded incitement policy be implemented or be operationalized?

Yes. And I think in addition to that, asking for general guidance on how we should prepare for the events of the next day, since we knew that there was a high likelihood for violence, so, in addition to asking if we could take action on some of the content we were seeing, asking for clarity and guidance around, you know, what should we be doing to prepare, how should we be staffing the next day, what should our plans
be, just looking for additional guidance for January 6th.

Q So, in your experience, was it unusual that your supervisors had not given you and your team these kind of marching orders for January 6th? For other events that had a similar potential for violence, whether in the U.S. or elsewhere, was it common practice for Twitter to have more concrete staffing orders and marching orders?

A It was not uncommon for safety policy during my tenure given some of the things that we have discussed, being the lack of leadership, the general disorganized nature of the team, but I believe that for -- I guess if you’re asking should there have been more guidance and guardrails, yes, there should have been. But up until this point, leadership had never been put together in terms of providing that guidance for an event like January 6th.

Q So, for example, for Election Day, were there guidelines in place for your team on how to be staffed, what the sort of triage process looked like in the event of violence occurring, a kind of special event war footing that you were expecting or hoping for on January 6th?

A Yes. On the Election Day, there were additional lines of communication where we could discuss things that we were seeing that was established by colleagues on my team, not by leadership. I helped create some of those channels for communication. And yes, it would have been helpful to have a sense of whether we could mimic some of those channels that we had for the election for January 6th, you know, if we were going to be taking that extra step of removing content, if we thought that the situation was escalating, what did that look like, needing that sort of guidance from leadership that was lacking.

Q So what was the response from your supervisor in Dublin when your colleague made these claims of concern?
A Our supervisor said that we should evaluate content under every other policy before evaluating under coded incitement to violence. And there were instructions given that we really should not be looking at coded incitement to violence unless we see some sort of violence break out on that day. And I recall members on my team asking for clarification, saying, Does that mean that we can’t take content down unless someone gets shot? What sort of violence needs to occur before we can take that extra step to take content down?

Q So your instructions the day before January 6th was that you can only use this policy, a policy that’s designed to prevent offline harm, once there was offline harm?

A That’s correct, yes.

Q And was it your impression that the supervisor in charge of this meeting was well-attuned to the possibility of violence on January 6th?

A No. I don’t think leadership in general, be it in Dublin or the United States, was taking the risk of violence seriously. I think there was the notion that members of our team who continued to raise this issue were overreacting.

Q And so, how did the conversation continue from there when the supervisor in Dublin essentially rebuffed your colleague’s request for additional guidance?

A So our supervisor said that we should evaluate content under all of our other policies, and if we get something under coded incitement for violence, we would need signoff from the recently hired supervisor in the United States who had been a part of the team for 2 weeks at the time and had not yet gone through training on our company’s policies.

Q So the same state of play, essentially that there was no clarity on how and when to apply coded incitement remained in place by the time this meeting ended?

A Yes, that’s correct.
Q And what was your sense of the danger that awaited when you left the meeting?

A I mean, it felt like there was a storm rolling in and we had no way to stop it. And we had tried our best to try and figure out how to navigate that day, and it was kind of an instance of we were anticipating the worse, but we couldn't even prepare for it because we didn't have the authority to take these tweets down. Despite asking leadership numerous times for guidance, for approval, we weren't getting it. So there was a sense of doom, in a sense, where it was one of those moments where we felt as though -- or I felt as though we tried as best as we could and whatever was going to happen on January 6th was going to happen, and we would just need to kind of meet the day as it came.

Q So just to summarize, you left this meeting, and there was no sense of emergency preparedness at the leadership level of Twitter?

A Yeah, that's correct.

Q And there was a final reminder that you did not have permission to freely utilize the coded incitement policy?

A Yes, that's correct. We were told that we needed approval again from a supervisor who had just been hired.

Q But, again, that sense of approval would not have been very helpful in a quickly moving situation on the ground, especially because you were told you could not even consider use of this policy until there was violence, correct?

A Yes. Yes, that's correct.

Q And you were not told precisely what that threshold was for violence that would allow you to use the coded incitement policy, correct?

A Correct. It was not clear. Again, a colleague asked, What does that
mean? Does someone need to get shot before we can use this? And the response to that was unclear.

Q And this is the same policy that had been ready to go for 2 months at this point?

A Yes, that's correct.

Q And it was your view that Twitter leadership and the leadership on that call were not expecting or prepared for violence to occur the next day, despite your warnings?

A Yes, that's right.

Q So I want to move on to the day of January 6th itself, and, obviously, all of its followers. If you could just briefly describe what the role was going to be as a member of the safety policy team when you woke up on January 6th.

A Yes. Going into the day, our role was largely the same as it was any other day, dealing with escalations, pieces of content coming to our team for us to evaluate whether or not it was violative, and working on other projects before the insurrection began. And once the insurrection began, one thing that --

Q Well --

A Sorry.

Q Oh, we can kind of take it, like, one at a time, because I think it might be helpful to understand what the posture was like for your team before the insurrection began.

You mentioned that you were working on other projects. So was it not the intent that the entire safety policy team would be crashing on monitoring the protest in Washington, that was not what was going on?

A That's correct. It was the beginning of the year. People were planning
what projects they were going to work on for that quarter. There were meetings early that day that occurred that had nothing to do with what was going on in Washington.

And, again, the message we got from leadership was it's business as usual, deal with escalations as they come, and go on a case-by-case basis. And, again, our team is not the operational side, so we are not out there day in and day out proactively reviewing content on the platform. We are evaluating content that is high-profile, gray-area edge cases that come to us whenever someone needs additional guidance from the policy team.

So as the day began, before the insurrection started, it was very much just business as usual.

Q Were you surprised that it was business -- I guess you wouldn't have been surprised because you had just been told the night before, but were you disappointed that there was not more of a concerted task force that was devoted to monitoring what was going on in D.C.?

A Yes. And I was particularly frustrated with the lack of communication from leadership, the lack of communication amongst our team. There was -- you know, as mentioned, it would have been really helpful to have a plan saying, Well, we should be monitoring what's going on on the platform so we can help inform the operations team. But, instead, they were saying, Well, what project should you be working on for Q1?

Q So at what point did you stop those meetings? Do you remember, like, what the continuum of that was when the safety policy team peeled away?

A Yeah. It was not until the insurrection really began.

Q So --

A Um --

Q Go ahead.
A I would say that once we saw that people were making their way to the Capitol, that was when I think that folks started paying more attention closely to the news, what was happening. But it’s important to note that at that time we got no indication from leadership to change course, to start monitoring trends on platform, to start standing up some of our emergency remediations, some guidance. Again, it was, from my recollection, just eerily quiet in terms of communication. Our team channels were silent.

Q How many people would you say were looking at January 6th-related traffic on your team the morning of the event?

A Yeah. It’s hard to say an exact answer. I think that some people might have taken it upon themselves to look at content, but I can only speak for myself, and I was not proactively looking at content until the insurrection began.

Q Got it. That’s really helpful to understand. So you were not seeing what the platform looks like in the morning, let’s say, before the President’s speech was really underway?

A That’s correct, yes.

Q Who was in charge that day of the safety policy operation?

A It was unclear. The supervisor at the time, as mentioned, it was -- they were in and out all week, and so it was unclear whether our supervisor was online and working, or whether we should have another point of contact. As mentioned in the meeting on January 5th, we were told to ask questions to our new hire. So, by and large, it was very unclear that day who was running the show. It was disorganized and, as mentioned, our team chats were silent. There was not a lot -- not a lot of movement going on.

Q So zooming ahead towards around midday, early afternoon, once people are
starting to move towards the Capitol, you abandoned your other meetings, I assume, and
you started to really monitor what’s going on on the platform. Is that correct?

A  Yes.

Q  So I was only going to ask, can you give us a sense of what happened when
you logged on to look what was going on? What were you seeing?

A  Yes. So to clarify, I think what I did that day was not anything directed by
leadership, at least in that period of the day, when the insurrection was beginning. So a
lot of what I was doing was following the news, and I guess in disbelief, and looking at
content on platform and asking in our team channels what should we be doing, should we
be starting to create guidance, asking questions from leadership. And during that time,
in regards to content I was seeing on platform, to be honest, I don't remember specifics.
I think it was a lot of the same, but it was just at that point, as mentioned, more specific
in terms of people saying that they’re going to the Capitol, that sort of thing. But my
memory is not perfect on all of the content that I saw that day.

Q  Sure. So two questions there. The first is, you said you were kind of
asking other members of your team, leadership, what you should be doing. What
responses were you getting when you sent these messages?

A  I wasn't getting a response. There was very little to no communication
from leadership.

At a later point there was a team meeting to kind of divvy up the work, but at that
point it was -- there was no communication. It was -- there was no response to my
questions.

Q  So we’ll get to that in a moment, but before we do, I wanted to ask more
about the content you were seeing. You mentioned people saying they were going to
the Capitol.
Do you recall whether that was in response to the President’s -- President Trump’s speech where he told folks to go to the Capitol?

A Yes. Yes, it was.

Q And how about content related to Vice President Pence? Were you seeing more than usual? Had there been a shift that day?

A I don't recall specifics relating to Mike Pence on that day.

Q Was there any discussion of firearms that you remember that morning, or other weapons?

A Yeah. If memory serves, I believe that there were tweets that I saw that were discussing that people there were armed, that -- just this, again, general tone of these people are -- they are ready to be violent. It was less theoretical at this point, and it was more they are armed, they are ready to take people down. So I do remember that type of content, yes. Any additional specifics to firearms, I am not sure of.

Q Well, that type of content which you just mentioned, was that something you were able to remove or it had to remain on platform because of coded incitement?

A Yeah. It would depend upon the specific tweet. So if there was -- again, if there was coded incitement of violence, like, We are locked and loaded and ready, that would not be taken down. If it was something saying, I have a gun and I am ready to shoot, insert target, then we would be able to take action on that.

So it's also worth noting at this point we did not have formal monitoring from our operations team on the platform. That happened later in the day.

Q When you say that formal monitoring by operations teams, were there other events of potential unrest around the world?

A Uh-huh.

Q Would there have been an opportunity or an option for it to assign its
operations team to monitor that kind of event?

A Yes, there would have been. So, for example, there were examples of other high-risk events, like the anniversary of the Christchurch shooting, where we might start to see violative content where we could proactively issue guidance to our operations team with key words and hashtags to serve as content. This was not something that we did for preparing for January 6th.

Q Did you or others on the safety policy team make a request for that sort of operation to be put in place?

A I can't remember specifically about guidance. I believe that we did ask whether we should be sending out some sort of monitoring guidance to our operations team. I can't remember specifics.

Q But nothing was done?

A Nothing was done.

Q And instead, the violence -- the guidance that stayed in place was this post-election guidance, a lot of which was copy and pasted from the summer, correct?

A That's correct, yes.

Q So some of these hashtags that we looked at earlier have to do with BLM and antifa and sort of burning of commercial property, correct?

A Yes, yes.

Q Which would have been wholly unhelpful to a moderation team looking at content that was about hanging Mike Pence or storming the Capitol?

A Yes. As mentioned, that was largely a copy-and-paste from summer of 2020, and it would not have applied or been useful for our operations team to surface violative content.

Q I just want to actually drill down on this point.
It was typical practice for Twitter to release policy guidance before a major high-risk event to help its moderation team look for hashtags and key words that would be associated with potential offline harm, right?

A That's correct, yes.

Q And that is, in a sense, what was done with the post-election guidance in early November, right?

A Yes.

Q But many of those terms and points of guidance were just copy and pasted had from what Twitter had already issued in May, June of 2020 related to the BLM protest and counter protests, right?

A That's correct. And the post-election guidance was supposed to have informed operation teams how to treat content that they come across and see if escalated to them as opposed to, for example, the guidance that we ultimately shipped on January 6th, or I gave the example of the anniversary of the Christchurch shooting, that is guidance for the operations team to proactively surface and look for violative content as opposed to guidance on what they should do should content be escalated to them.

Q Got it.

So there was never any proactive guidance issued, it seems like, in relation to the post-election period at all?

A Yeah, that's correct.

Q But, in particular, there was no proactive guidance issued related to January 6th?

A That's correct.

Q Would your team have been in charge of helping to put together that
We would have, yes.

Would it have been easy for you to establish a list of hashtags and key words that would have been associated with potential violence at the Capitol?

Yes. We had been seeing these trends for months, so I think that we should have easily pulled together key words and hashtags. Obviously, there would be certain elements that evolved the day of. I think that some of the content involving Mike Pence we might not have been able to anticipate. But, by and large, I think we could have put together a nearly comprehensive list of hashtags and key words in advance.

Things like Stop the Steal, Civil War 2.0, locked and loaded, that kind of thing?

Yes, that's correct.

But you were never asked to create that kind of document?

Not prior to January 6th.

And on the day of, you were asked to create that kind of document?

Eventually, yes. So once the Capitol had been attacked and there were insurrectionists inside the Capitol, there were a series of meetings that occurred with select people from health policy, which includes safety policy, insight, integrity. I did not attend those meetings. But afterwards we had a meeting with just members of safety policy, and the direction was, we needed to produce two work streams.

The first was assessing Donald Trump’s tweets and essentially what they were going to do with his account on that day, which a colleague of mine was responsible for producing.

The second work stream was guidance for our operations team to -- how they
should action certain content that they are seeing, and that was the proactive guidance
that we were asked to ship. So at that point, other colleagues and myself worked to
create guidance that we did ship to our operations team for how to handle content they
were seeing that day.

Q    Got it.
So this was around -- the Capitol was breached around 2:00 p.m., 2:15 p.m., so
you would say that's approximately when you had these meetings?

A    Yes.
Q    Okay. And you were, at that point, told to figure out what to do with the
President's account and what to do with content that was encouraging the ongoing
attack?

A    That's correct, yes.
Q    So I want to get into those two work streams, as you called it, but first I had
some questions about what Twitter has told the select committee about January 6th
itself.

So, if we can pull back up the exhibit 3, which is the timeline.
And so at the bottom of page 17 on the timeline, there's an entry at 2:24 p.m.
where President Trump tweets an attack on Mike Pence, saying that he didn't have the
courage to do what should have been done to protect our country.

Were you aware of this tweet eliciting a reaction from Twitter leadership?

A    Not at this point. As mentioned, there were some meetings that were
happening behind closed doors that I wasn't a part of. And at this point I was more
concerned with getting that guidance put together and seeing what was happening on
platform.

Q    Okay. And just one last question on this.
Right underneath that -- you don’t have to scroll, Heather.

But right underneath that, 4 minutes later, that tweet has a label applied to it, an interstitial. Is that 4-minute gap unusual, given what we talked about with Twitter’s reluctance to act on President Trump’s account?

A Yes, yes, very unusual.

Q What is your reaction to seeing that 4-minute gap right here on the document?

A I think that -- well, one, I can’t necessarily remember all of the details to comment on the veracity of that time gap, but knowing that if there were 4 minutes between the President of the United States tweeting something and a label on it, that was highly unusual. And I think it just goes to show the fact that, I don’t know, maybe at that point Twitter was starting to understand the gravity of the situation.

Q You can see it actually in other -- almost an hour and a half before Twitter decides to restrict engagements with the tweet.

A Yes.

Q Given what was happening at 2:24, where there was an ongoing breach of the Capitol with people calling for the Vice President to be hanged, why do you think restrictions on that tweet weren't put in place 4 minutes later when it was first labeled?

A I think that this is another case of Twitter trying to walk this line of showing that they're doing something, but not doing enough. So they label the tweet, but that label doesn't do much of anything, and then it takes further deliberation and time to see the situation on the ground evolve and escalate before they take further action. And so, I think it’s this continuous notion of Twitter, like, in a way, trying to keep both sides happy and not just applying for policies consistently and doing what they think is the right thing to do.
Q Understood.

So was it your view that Twitter had a comprehensive strategy towards how it
would treat President Trump’s account as the events unfolded, or it seems a little more
reactive based on what you’re saying?

A It was definitely more reactive. I don’t think that there was a plan in place
for how to treat Trump’s accounts ever, and it took an insurrection for them to ultimately
suspend his account.

Q So I want to scroll up just a little bit, if we could, Just a little bit further -- oh, wait. Actually, no, we’re all good. Sorry. Down
again.

You can see at 2:20 p.m., Twitter begins to identify and remove -- sorry -- and
review content or coded language that has the potential to incite violence.

This seems like perhaps the point at which you were told to figure out how to deal
with the content that was on the platform.

Would that be accurate?

A Yes. I think that one thing that is still missing is the remove aspect. So we
could identify and review content with coded incitement to violence, but there was still
that missing aspect of actually being able to remove it from the platform.

Q So you’re saying that there was no cohesive guidance at 2:20 p.m. that you
were able to tack on coded language even at this point?

A Correct. There was never, to my knowledge, leadership convening a
meeting and saying, Violence has broken out. You have the green light. Take it all
down. That never happened.

Q Was there a point at which you or others on the safety policy team started
taking some tweets down anyway?
A: I can only speak for myself, and I was not taking content down. I was looking for examples of content to include in the guidance for our operations team to give them examples of what to look for when they’re doing their proactive monitoring.

Q: So even then, as there are rioters approaching the House chamber and rioters in the Senate Chamber, you still only feel empowered to look at tweets to include a proposed guidance, and not to remove them?

A: That’s correct, yes.

Q: And this was because you have been told to create a proposal that had essentially been in the hopper and ready to go for months at this point?

A: Yes.

Q: So if we could scroll down to the next page, to the 5:00 p.m. entry. You can see the final entry on the page that around 5:00 p.m. Eastern, "Safety policy circulates additional guidance to minimize content with the propensity to incite violence."

Would this be the document that you were tasked with creating in the meeting in the early afternoon?
[3:40 p.m.]

A Yes, that's correct. This would be the guidance I worked on.

Q And so you were part of the team that was involved in drafting this document?

A Yes, that's correct.

Q Which essentially was the culmination of the coded incitement policy combined with some proactive guidance on how to deal with content related to January 6th in particular?

A Yes. Now, one thing that I will note is that with coded incitement to violence specifically, I remember not including it in the guidance that went out to our larger swap of operations teams. That policy was written only for safety policy, and there was -- again, leadership reiterated that it should be used as an absolute last resort. So it's -- my memory is such that this guidance was more our run-of-the-mill policies and how to apply them and not going above and beyond to take extraordinary action given the situation.

Q You mean at that point you were not sure that you would be allowed to get that policy through quickly?

A Yes, that's correct.

Q I want to actually take a look at that document in moment. But, first, while we have this up, if we could look at the top of page 18. I believe it's -- it's there. So it's the last entry that you can see on the screen here, that Twitter reinstates certain interventions originally established in October 2020, and that will just continue after the election to disable engagements on certain tweets. Do you know what this is talking
about?

A No, I do not.

Q Were you aware of certain precautions that were put in place about disabling engagement on certain tweets with information or potential violence that were discontinued after the election?

A No. From the safety policy side of things, which, again, would cover wishes of harm, et cetera, we were not told that there were going to be particular interventions added on January 6th, like disabling engagements. So whatever this means, I do not believe that it is pertaining to safety policies.

Q Okay. Got it. So at this point, can we pull up exhibit 4? So is this the document that was released by the safety policy team and shipped out?

A It looks like it, yes.

Q Do you want to take a few minutes, and we can go off the record, and you can review it?

A Yes, that would be great.

Q Great. All right. So we’re off the record at 3:44.

[Discussion off the record.]

Q All right. We can go back on the record at 3:46.

BY:

Q So having reviewed this document, again, I wanted to ask you about some of the language regarding coded incitement in the document, as well as sort of your recollection of creating this document, things that you’re able to talk about.

A Yes. So one of the pieces that I think was more focusing on it, obviously the coded incitement to violence section. And you’ll note that the instructions say to
escalate to the safety policy. So what this means is that, while we might have received an uptick in escalations and we might have more content surfaced, we are still running into the same problem where that content needs to be approved at our level from one of the safety policy supervisors before we could take action. So though I think that this guidance is helpful in surfacing that content, again, it is not achieving the goal of actually removing it from the platform. And as we know, leadership on safety policy was reluctant to delete content under this policy.

Q So this document helps to ensure that safety policy team members are aware of inciteful content on the platform but still doesn't guarantee that it's removed?
A That's correct, yes.
Q And that's because of the existing roadblock in Twitter's leadership structure regarding gaps and enforcement on incitement, correct?
A That's correct, yes.
Q And just to illustrate this, this means that if someone at 5:30 p.m. Eastern has folks that are being cleared out of the Capitol by the National Guard came across a locked and loaded revolution 2021 tweet, they would be able to and should show it to you and your colleagues from safety policy, but they would not be empowered to remove it?
A That's correct, yes. And to underscore the issue with that is the time that it takes for that tweet to be posted, for it to be surfaced and found by our global escalation team, otherwise known as GET, forget to upload it to safety policy, and safety policy to escalate it to our leadership, the time where that tweet can do harm is likely already done. And we also know that leadership on our end isn't going to want to remove it at the end of the day anyways.

So to say it was a superfluous element to the guidance might be a little too far,
because potentially there could have been an instance where leadership said, yes, take it down, but this was not making the content on the platform safer or healthier.

Q And you were still not receiving clear signals from your supervisor about what they believed to be the threshold for removing content, correct?

A Correct.

Q Were you receiving information from anyone else above your supervisor about what you should be removing when it comes to potential incitement?

A No. There was no clear guidance. And by the time that I left the company, there was still no clear guidance on coded incitement to violence. It was a policy that I think after January 6th no one wanted to touch because it just was incredibly frustrating and difficult to get the answers around how we should treat that type of content.

Q So I do want to get to the after January 6th developments soon, I hope. But I did just have a couple more questions on the day and the day after, 2 days after. So shifting away from the feeling -- the attempts to deal with coded incitement and content related to the attack, I wanted to get back to the question about President Trump's account. So at the end of the day, 5, 6 p.m., as the Capitol is being cleared and you sent out this policy, what were your expectations for what would happen with President Trump's account? He obviously continued to tweet that day.

A Right. I personally did not know what was going to happen. As mentioned, I was aware that a colleague of mine was working on an analysis of his tweets and his account. At that point, the trend that I had seen was a lack of action on Trump's account. So I did not necessarily anticipate that leadership would change course, despite the unprecedented events of the day.

The decision to block Trump from posting for 12 hours was announced to the
broader safety policy team, 5 to 10 minutes before it was tweeted publicly, which, again, highlights the lack of communication amongst the team that day. It was extremely unclear how we came to that determination and the rationale that was used to get there.

Q And was it your sense that there was a game plan what would happen after the President's time out expired?

A There was no game plan. It was another instance where I would have appreciated leadership to put some plan in place for evaluating his tweets, given we publicly stated if he violates our policies again, we would permanently suspend him from the platform. So I knew that there would be intense scrutiny over every tweet and internal and external pressure to understand what our analysis on each tweet was. There's no plan in place. And at that point, I was unsure if we would eventually take that step to permanently suspend his account. It was one of those instances where I just couldn't predict what leadership would decide.

Q And before -- I want to move on to January 7th and January 8th, but my last question on the 6th itself, was the impact of the mismanagement of your team on how you were operating that day, how you were surfacing tweets and potentially removing tweets, even beyond the coded incitement, just the full on violent incitement, what did you see as the barriers to a more effective operation?

A Communication, I think, is a big part of that. There was, as mentioned, little to no communication throughout that day, besides the one meeting where it seemed like decisions had already been made in regards to what work products needed to be produced. And I think the inability to have decisions being made, it would have been helpful, I think, to have said there's clearly violence happening. We need to remove this content under coded incitement to violence, full stop, and we never got that. So that was definitely a barrier as well.
And I think some of the -- the disparate conversations that were happening and the meetings behind closed doors made it very difficult to effectively do my job and understand how I could use my time most effectively that day.

Q Do you think that made a difference in what ended up staying on the platform throughout the day?

A Yes. Yes. I think that we should have enforced more content that day than we did.

Q And apologies, I know I said last question. But can you give a sense, once you were locked in and looking at the content in the midafternoon, what were you seeing as people moved through the Capitol?

A There was, I think, a lot of, as mentioned, there was hashtag commentary about revolutions, Civil War 2.0, when the woman who got shot, whose name I am forgetting --

Q Ashli Babbit.

A -- Ashli Babbit. There was a lot of discourse around that, around potential graphic images of that that we had to either take down or put a label over because it was instant media. So I think that there was a lot of content that I was seeing that was almost more reacting to the situation than at that point the insurrectionists themselves posting content. So more just commentary and shock at that point.

Q Did you see a discussion of specific breach points? Like, come to this part of the Senate, come to this part of the House, conflict with law enforcement here, that kind of thing?

A It was discussion around what had been breached and where people were trying to get to and certain offices that people were heading to. I don’t remember specifics if there was a call to action in that in regards to come here, but there was
absolutely discussion around specific locations in the Capitol where things were happening.

Q And were those tweets able to be removed under any policy?
A No.

Q Would they have been removed under a coded incitement?
A It depends on how it was phrased. But, likely, it's hard to speculate on hypotheticals.

Q No, of course. That's helpful.

All right. Well, I wanted to move now to a few questions about what happened after we were discussing President Trump’s account and its future. So can you add your involvement in the debate and deliberations around President Trump’s account on January 7th and January 8th?

A Yes. On January 7th when Trump was able to tweet again, it was in the evening, and that tweet, the team reviewed collectively and assessed that it was not in violation of our policies. Pretty straightforward. The next morning was when Trump tweeted about 75 million American patriots. And the timing of our assessment was such that it was when our colleagues in Dublin were about to sign off and we call that a handover. So at that point they had done an initial assessment and determined it was not in violation of our policies, but had not done a full -- gone through the approval processes in the sense of getting signoff from leadership that that is indeed the case.

They handed it over to myself and my other colleagues in the United States. And we looked at the assessment again, and we thought that it was in violation of our glorification of violence policy. So we started an assessment that details why we thought it was in violation of our policy. And one of the pieces of evidence that I hope to contribute to this assessment was the response to his tweet that we saw from his
supporters, the way that it was being interpreted on the platform. And it was such that
people were interpreting it as, you know, the fight’s not over; they’re going to keep going;
that’s what Trump wants us to do; he wants us to not stop. And, you know, we had a
sense of other dates that were coming up, be it the inauguration, be it later in January.

So that was the integral part of our assessment in regards to why we found this
tweet as violent and ultimately leading to his permanent suspension.

Q So if I’m understanding correctly, you were concerned by the fact that
President Trump supporters on Twitter seemed to be interpreting this tweet, the
[inaudible] among the great American patriots, as a kind of endorsement of what had
happened on the 6th and a endorsement of potential violent efforts in the coming days?

A Yes.

Q And that was, therefore, a violation of your glorification of violence policy?

A Yes.

Q And can you just explain briefly what the glorification of violence policy,
which you talked about before, was meant to do?

A The policy, historically, was applied to praising violence, be it mass murders,
shootings, so it was head profile violence events, and condoning, praising, glorifying,
endorsing that type of action.

Q So the response to those kinds of tweets -- praising mass casualty events -- I
assume it’s part of the analysis when you’re looking at any glorification of violence and
whether it should be removed is part people on the platform looking at it that way, right?

A Yes.

Q So the responses to President Trump’s tweets fit that pattern?

A Correct. And one thing that I will note is that the response to my tweet is
going to solicit potentially less of a reaction from other users than Donald Trump. So
that's not an ability that we have every time we're analyzing a tweet against glorification of violence, right? But in this instance where we could see how his supporters were interpreting this, it became clearer what it meant.

Q So your concern was really that his tweet would be used as a groundwork to justify future acts of violence?

A Yes.

Q And was the caliber of content in response to that tweet similar to what you had been seeing, for example, in response to his December 19th tweet?

A Do you mean was our action the same from December 19th?

Q No, no. I mean, were you seeing his supporters on Twitter respond in a similar way as you had seen them respond in the buildup to January 6th?

A Yes. Yes, that's correct. It was an indication that it felt like January 6th wasn't the end, that Trump was endorsing the continuation of the efforts of the insurrectionists.

Q And I believe there was a hashtag associated with the replies to that tweet that was #J17. Is that correct?

A Yes.

Q And can you explain why that was of concern to you and your team?

A Yes. That was another date that we thought could be a January 6th 2.0, another instance where violence could break out. We know that, you know, there were plans for rallies, protests, whatever you want to call them at other State capitols. And we thought that January 17th was going to be another date where violence could break out.

Q So you were against seeing a specific date targeted by supporters of President Trump in response to his tweets that seemed to implicitly endorse that action?
A Yes, that's correct.
Q Did it give you a sort of sense of déjà vu?
A Yes. Yes, most definitely. And, again, it just felt like there was a potential for January 6th to not -- not be the end of it, for it to continue, and for the momentum to continue to increase.
Q Did you -- can you give us a sense of scale? Are we talking about a level of replies to President Trump engagement with the tweet that suggested that future events could be potentially worse than what you saw on January 6th?
A Yeah, I can't speak to specifics in regards to the number of replies. But anecdotally what we did see was that his base, from the replies that I saw, were emboldened by this and by what he was saying. And it did seem like a continuation of everything we had seen leading up to January 6th. And we know that, you know, at that point, his supporters did what they did by and large because of what Trump had been saying for months. And Trump was continuing with that same rhetoric, so we presumed that the same pattern of behavior would continue.
Q Understood. So an hour after that tweet about the 75 million patriots, President Trump tweets again saying that, to those who are asking, I'm not going to be attending the inauguration. And that's his final tweet on the platform.
Did the reply to that tweet raise similar concerns about potential implicit endorsement of violence?
A I can't remember specifically. I think at that point we were heads-down working on the assessment for that first tweet. So I don't quite remember the replies to that second tweet.
Q That's helpful. So your suggestion -- your team suggestion that President Trump be permanently banned from the platform really was focused on that 75 million
That’s correct, yes.

Okay. And so you drew up this assessment, correct?

Yes.

And at some point, did you have any meetings handing off the assessment to your superiors?

Yes. There were a handful of meetings that I was not a part of, but it did make its way up the ranks, and finally to Vijaya and Jack Dorsey to sign off on.

You were not involved in any meetings with, for example, Del Harvey to discuss the evaluation, the assessment?

No. A colleague of mine was, I was not.

Okay. So your last firsthand engagement with the debate over President Trump's account was finishing that memo and then it was out of your hands?

Yes, that’s correct.

Okay. Thank you. So, in that case, I just wanted to ask like one or two questions about the broader implication for what this whole episode meant. We talked about how Twitter treated President Trump’s account, and I wanted to ask you how that reflected in this final saga of his account.

For example, would any other user have still been on the platform for 8 hours after the responses you received, his 75 million patriots tweet?

No. I think that there was a lot of time taken to deliberate the permanent suspension of the U.S. President. I think that if it was a regular user, I don’t think that Trump would’ve gotten to that point. I think that if it was a regular user, he would have been suspended much earlier.
Q I know that your sense that Trump's account was in danger of suspension -- excuse me -- would have been in danger of permanent suspension had your team not taken the time to point out these ongoing threats of violence.

Ms. Ronickher. Can you ask that again? I'm sorry, I'm not sure I followed.

Sure, sure.

Q So you mentioned that Twitter said publicly that if there was another violation of its policies, the President would be suspended permanently from the platform. And then he comes back on the platform, arguably violates the policy with his first tweet on January 8th, but your team went through and collected evidence about why there were replies to that tweet that showed an ongoing sense of harm. So I guess my question was, was that ongoing harm a necessary piece of his suspension?

A It's hard to stay exactly whether -- whether or not if we didn't include that, whether the assessment would have still been approved by Jack Dorsey. That being said, I do think that it was compelling evidence to understand how it was being interpreted by his supporters. So I can't necessarily say one way or the other what the outcome would have been, but I do think that it was an important piece of information that was included in the assessment.

Q That's real helpful. Well, I think that's all I have on the suspension debate. So I did want to ask a few questions on what happened sort of longer term after January 6th, follow up with some concluding questions, but I don't think that will take very long. And I recommend we take a 15-minute break before and come back at 4:30. That works for you both?

The Witness. Yeah.

Thank you. So we can go off the record at 4:14, and I'll see you all
soon.

[Recess.]

Okay. So we can go back on the record at 4:30 p.m.

Q At this point I wanted to ask whether you were aware, after January 6th, whether Twitter kept the coded incitement policy in place that was sort of part of that guidance bundle that you sent out on January 6th, or if they took any other steps to fully implement the policy.

A No, they didn't take additional steps to further implement the policy. As mentioned, by the time I left the company, there were still questions around coded incitement to violence and how or if it could be applied to other situations, such as the Brazil election or the U.S. 2022 midterms.

Q Did you believe that this policy was still necessary?

A Yes. Absolutely. I think that there were going to be additional instances where we would want -- Twitter would want to use that policy. And so it would have been helpful to have clear guidance on it moving forward than what actually had been used.

Q Do you recall any specific conversations with your supervisors in which you urged to them to take steps to codify the coded incitement policy after January 6th?

A Yes. I was aware of preparations for the Brazil election that is happening next month, and there were conversations around the likelihood of similar rhetoric spreading on the platform and questions around whether or not we could use coded incitement to violence. So I raised that to my supervisor, with no clear answer in response.

Q And this is rhetoric similar to what we saw around President Trump's tweets
leading up to January 6th?

A Yes. One of the phrases I became aware of in the Brazil market was "make justice with your own hands."

Q So now more broadly, the failure to have a coded incitement policy is not only affecting Twitter's ability to respond to events in the U.S. but outside of the U.S. as well?

A That's correct, yes.

Q So it's our understanding that the safety policy team repeatedly requested that a retrospective be conducted about Twitter's actions surrounding January 6th and the suspension of President Trump. Is that true?

A Yes, that is true. There were numerous asks, roughly about every month, for the safety policy team specifically to conduct a retrospective around our team's involvement leading up to January 6th and the permanent suspension of Donald Trump, given the policies we dealt with were the policies that were largely involved that day and ultimately led to Donald Trump's suspension.

Q So to your knowledge, was there ever a retrospective conducted?

A By the time I left the company, no retrospective had been conducted.

Q Was that unusual for an event of that scale?

A Granted, I did not have experience with similar events of that scale during my tenure at the company, but there was a retrospective done during my time at the company for an event that, in my opinion, was significantly smaller than the insurrection that happened on January 6th. So I believe it is unusual, yes.

Q Would you be able to tell us what that event was?

A I don't remember the specifics. It did not have to do with content in the United States. I can't -- I can't remember specifics of that case.
Q That's all right. But if you could recall the response that you got when you requested a retrospective, what was -- what was leadership and your supervisor saying about why they weren't creating that kind of document?

A I was told that it was not a priority for the team at that point.

Q And why did you think it should be a priority?

A Because there was an insurrection on the U.S. Capitol, and we suspended a sitting U.S. President from the platform that he liked to use the most. And I think that it's important to understand what part our team played, how we could have been better prepared. That was probably one of the biggest, if not the biggest decision in the company's history. And I think it's important to analyze it and understand what went well, what could have gone better, everything in between.

Q So why do you think that Twitter leadership was hesitant to make this kind of document?

A I think in many ways leadership was exhausted with focusing on the U.S. election. As I noted previously, leadership expressed frustration when colleagues, myself included, would asked about our North Star, would ask about guidance for coded incitement. And I think that they wanted to put the U.S. 2020 election behind them. And there's another part of me that thinks there would have been a lot of -- a lot of bullet points in the where-we-could-have-improved column. And I think that leadership was averse to hearing that feedback.

Q Got it. So the select committee has discussed the possibility of these retrospectives with the current global lead for safety policy at Twitter, Juan Felipe Rincon, who said he hasn't really reviewed Twitter's actions on January 6th. And he said that only folks who need to know that history need to review them.

Would you consider Mr. Rincon to be someone who is in the need-to-know
category for reviewing documents about Twitter's response to January 6th?

A Absolutely, he is someone who needs to know.

Q Why do you think it's important for the global lead for safety policy to be familiar with Twitter's response to January 6th?

A As the new lead for site policy, he will be in charge and responsible for making decisions relating to upcoming elections or potential civil unrest, be it in the United States or globally. And there was a lot of learning from what happened to January 6th. And it's important to be familiar with what happened to try and prevent that from happening again.

Q Thank you. And my apologies; he is the global lead for site policy, not safety policy. So what do you make of this claim that he's not in this need-to-know category? Is that concerning to you?

A Yes, it's very concerning and, I think, naïve to believe that something like this won't happen again.

Q Did you ever discuss January 6th and Twitter's response to the attack on the Capitol with Mr. Rincon while you were at Twitter?

A Not specifically, no, I did not. Our overlap was limited by the company.

Q While you were overlapped, did you ever hear Mr. Rincon discussing January 6th or sort of the need to prioritize certain issues related to code incitement or violence on the platform, that sort of thing?

A He mentioned that we needed to audit our policies and improve some of them, including glorification of violence, the policy Trump was suspended under. But specifically as it relates to January 6th or the suspension of Trump, he never said it in so many words tying it back to those two incidents.
Q So what do you mean he -- what was your understanding of what he meant by auditing those policies such as glorification of violence?

A I think he saw that there were, again, issues with the consistency in which we applied our policies. There was ambiguity that we needed to clarify and needed to go through and audit the policies to make sure that they were clear as they can be and you can apply them consistently.

Q Understood. And as a final note, I just wanted to ask, when the select committee spoke to Mr. Rincon about whether he reviewed documents related to January 6th, and we intimated that it might be important for someone to view those documents in this position, he implied that we would have simply different management styles. And I'm wondering, what kind of management style, in your view, would be necessary to not review those kinds of documents?

A The irresponsible management style, I would say. I think that it is -- it is negligence to not be aware of what happened leading up to January 6th and the suspension of Donald Trump. And as the leader of a team who is integral into making that decision, not being aware of what happened, to me, is -- it's not a matter of management style; it's a matter of being a responsible leader and making informed decisions.

Q And, obviously, there was another online discussion about the potential attack on the Capitol in March of 2021, and I believe in August as well, as well as well-publicized instances of far-right violence or attempts of violence in the United States since the attack.

So are you concerned, given what we’re talking about, about Twitter's ability to track and defuse some of those conversations happening on its platform?

A Absolutely. That's why I'm here, I think, is just to -- I think that in order to
address some of these concerns, it's important to understand what happened and have
an accurate reflection of the events leading up to and after the insurrection of
January 6th and the involvement that Twitter had. And it is concerning to me that there
is seemingly no appetite to reflect on that and learn from that, work towards making it
not happen again. I think that the people who plan these types of events do so in a
coordinated way where Twitter needs to work hard to make sure that something like this
doesn’t happen and not turn a blind eye towards it.

Q So since -- to the best of your knowledge, since January 6th, Twitter has not
implemented a policy that the safety team developed to prevent calls to violence from
proliferating on the platform?

A By the time I left the company, no, they had not.

Q And they have not conducted a retrospective, as far as you know?

A They had not.

Q And based on our conversation, it does not appear as though the global head
of the site policy who’s in charge of these responses intends to conduct such a review or
be familiar with it?

A Yes, that sounds correct.

Q So before we conclude, I wanted to ask a few questions about what that
means for the threats we are seeing today. And just a few weeks ago, we saw an FBI
office in Cleveland attacked by someone who it appeared was responding to President
Trump's tweet -- or posts on Truth Social about the FBI's raid of Mar-a-Lago. Are you
familiar with that instance?

A Yes, I am.

Q And does that remind you of the kind of call and response dynamic that
cornsers you in the aftermath of the 2020 election?
A Yes, absolutely. I think that gearing up towards U.S. midterms and eventually the 2024 Presidential election, I'm deeply concerned about what's going to happen and the role that social media is going to play and how actively they are going to try and mitigate the events of January 6th from happening again.

Q So talking about Twitter specifically, there are some posts that we were able to identify that I wanted to run by you.

Just hours after the raid on Mar-a-Lago and President Trump's tweets about it, Steven Crowder, who's a prominent conservative commentator posted on Twitter, quote, "Tomorrow is war. Sleep well." And as of this deposition, Mr. Crowder is still on Twitter. Does that concern you?

A Yes. That's the similar rhetoric to what we were hearing before the insurrection.

Q Oh, go ahead.

A You know, the codes used might be different, but, in my opinion, the intent and the meaning is the same.

Q So this would not suggest to you that Twitter has learned its lesson from what happened on January 6th?

A That's correct.

Q And in another tweet, Marjorie Taylor Greene, a Congresswoman, called the raid, quote, "The rogue behavior of communist countries." And said, quote, The types of things that happen in countries during Civil War. That was just hours after the raid.

So, again, we see a mention of Civil War and no action by Twitter. Is that something that you would be raising the alarms about if you were still working there?

A Yes, absolutely. I think that there is, again, the need to have clear guidance, clear understanding of the processes. And it sounds like, to date, that hasn't
changed. There is still ambiguity. Things are being left on the platform that are clearly problematic.

Q So mostly these messages are being shared on fringe platforms like Truth Social where President Trump is still active. Are you concerned about what happened if President Trump were able to rejoin Twitter?

A Exceedingly concerned about that, yes.

Q Why is it different for President Trump to be posting these things on Truth Social versus on Twitter.com?

A Truth Social I don't think has as big of a reach as Twitter does. And I think that his message will be amplified more on Twitter than it is on Truth Social.

Q And one final point on that line of questioning. President Trump also recently started sharing posts from Q. The posts are allegedly the center of the QAnon conspiracy theory on 4chan and 8kun. He was sharing some Q posts on Truth Social. Is that concerning to you?

A Yes, absolutely. I think that that's something that should be a black and white decision of not being allowed. But, yes, very concerning.

Q And so as someone who's been -- who worked in safety policy at Twitter for years, the fact the President of the United States is sharing -- former President of the United States is sharing QAnon posts, does that make you concerned about what kind of content will be moving, migrating towards Twitter and mainstream platforms?

A Yes, definitely. I think that it will be interesting to see an interplay between this and accordingly how a company like Twitter is going to respond to that.

Q So do you think Twitter is prepared for the continued lies of this kind of rhetoric in the months ahead?

A No. I obviously no longer work at the company. But unless there has
been a drastic change, I am concerned about their preparedness for these upcoming world events and the trends that we're seeing with the rhetoric from former President Trump.

Q Thank you. Just a few more questions.

Given our conversation here today, could you again summarize your feelings about Twitter's responsibility for what happened at the Capitol on January 6th?

A Yes. I think that Twitter and its leadership was grossly negligent in its handling of content leading up to and after January 6th. There were multiple people who raised the alarm about how to treat this content and why it was problematic, and it consistently got ignored.

I think that Twitter had an idea of its negligence after the fact, which is indicative by a multitude of things. One being its reluctance to conduct a retrospective. I got paid a spot bonus for my work during the insurrection, which was unusual and, quite frankly, uncomfortable. And I do not think that they are willing to learn from their shortcomings, and it will result in them not being prepared in the event that history repeats itself.

Q What made you uncomfortable about receiving a bonus for your work during January 6th?

A Because an insurrection happened, and the work that I was trying to do was trying to prevent that from happening. And it felt almost like, I don't know, dirty money.

Q So are you worried about the potential for another January 6th, whether in 2023 when a new Congress comes into session, or 2025, especially when there's a new Presidential election being certified?

A Yes, absolutely. Very concerned about that.

Q And what actions could your team have taken in the runup to January 2021,
or do you think a team should take in the runup to January 2025, that could have made a
difference?

A Yeah. I think that, for one, removing the content we were seeing under
coded incitement to violence. And I think that it’s difficult; content moderation is very
hard. And I don’t want to negate or minimize the complexities involved in that. That
being said, there were pretty clear indicators in the lead-up to January 6th that I think
would have made it potentially less likely to occur, that being the content we were seeing
from the former President saying that the election was rigged, repeating that over and
over again multiple times a day, and the response that was listening from his base. They
were ready to become violent in order to ensure that he remains President.

Q And we talked a lot today about the warnings that you and the safety policy
team gave to both your supervisor and Twitter leadership about the signals you were
seeing. So I just wanted to ask one more time, at the end of the day, why you think
those concerns, those fears were not taken seriously and you were not given the leeway
to do what you laid out just now.

A I think that Twitter didn’t want to be seen as biased and stifling the voice of
Trump’s face or Trump himself, and they were scared of the repercussions of doing so.

Q And final question. What’s your greatest fear for how Twitter will handle
the continued pattern of violent extremism and the nexus with former President Trump’s
speech going into the future?

A I think my greatest fear would be Trump coming back onto Twitter and
Twitter continuing to not take action on content that we know is harmful and can lead to
offline harm.

Q Thank you. I think that’s all the questions I have. I’m not sure if [redacted]
if you have any final thoughts.
1 Seeing none, I want to give Ms. Ronickher or the witness a chance to say
2 something before we go off the record.
3
4 Ms. Ronickher. Do you have any final thoughts you want to share?
5 The Witness. No.
6
7 All right. Well, in that case, the deposition will stand in recess,
8 subject to the call of the chair. And we can go off the record at 4:57.
9
10 [Whereupon, at 4:57 p.m., the deposition was adjourned, subject to the call of the
11 chair.]
Certificate of Deponent/Interviewee

I have read the foregoing ____ pages, which contain the correct transcript of the answers made by me to the questions therein recorded.

_________________________________

Witness Name

_________________________________

Date