

May 17, 2021

Chairman Frank Pallone, Jr.
Ranking Member Cathy McMorris Rodgers
U.S. House Committee on Energy and Commerce
Attn: Joe Orlando
2125 Rayburn House Office Building
Washington, D.C. 20515

Dear Chairman Pallone, Ranking Member McMorris Rodgers, and Members of the Subcommittees:

Thank you for your questions for the record from the March 25, 2021 hearing entitled “Disinformation Nation: Social Media’s Role in Promoting Extremism and Misinformation.” Per your request, attached are the answers for the record to your questions.

Sincerely,

Facebook, Inc.

cc: Chairman Mike Doyle
Subcommittee on Communications and Technology

Chairwoman Jan Schakowsky
Subcommittee on Consumer Protection and Commerce

Ranking Member Robert Latta
Subcommittee on Communications and Technology

Ranking Member Gus Bilirakis
Subcommittee on Consumer Protection and Commerce

The Honorable Jan D. Schakowsky

- 1. Please provide our committee with a list within 30 days of all the universities, think tanks, NGOs, publications, journalists, activists, and activist groups that Facebook currently supports financially or has provided more than \$50,000 in funding to in the past five years. This should include all grants or other payments paid through a third party that ultimately produced reports, studies, editorials, or other publicly available content.**

Facebook regularly partners with third-party experts, academics, and other stakeholders on a variety of initiatives. For example, through Facebook Research, from time to time we invite academics to propose research in specific areas that align with our mission of building community and bringing the world closer together. Invitations typically address specific research challenges that we have identified are important in our mission. These vary in nature and may solicit proposals from individual researchers addressing a well-defined problem or may require multiple university departments to collaborate with a diverse set of skills to solve a more complex multi-disciplinary challenge. Calls for proposals align to specific areas of research. They offer insight to the academic community on topical research questions of interest to Facebook, where academic insight is key. A list of Facebook Research award recipients dating back to 2015 is available here: <https://research.fb.com/research-award-recipients/?dateYear=2018>. Additionally, the most recent announcements first appear on our blog at: <https://research.fb.com/blog/>.

- 2. Does Facebook share data on disinformation with independent scholars or researchers? If so, what information is shared, with whom is it shared, and what is the process for sharing it?**

We regularly work with outside academics and researchers, including on projects that relate to disinformation. For example, earlier this year, we announced a partnership with a team of independent external academics to conduct objective and empirically grounded research on social media's impact on democracy. We want to better understand whether social media makes us more polarized as a society, or if it largely reflects the divisions that already exist; if it helps people become more informed about politics, or less; or if it affects people's attitudes towards government and democracy, including whether and how they vote. We hope that the insights these researchers develop will help advance society's understanding of the intersection of technology and democracy and help Facebook learn how we can better play our part.

- 3. Clarify who is responsible for Spanish-language content moderation in the U.S. Please indicate which specific team is responsible for that work and share the title of the manager who most directly carries that responsibility. In parallel, please also indicate the title of the manager who is responsible for efforts to counter Spanish-language disinformation in the U.S. and the name of the relevant team working on this issue. In both instances, please identify where these teams and managers sit in the organizational chart and to whom they report.**
 - a. Furthermore, clarify whether Spanish-language content in the U.S. is a specifically identified and assigned responsibility. Do the teams and managers involved have specific responsibility for Spanish-language content**

in the U.S. or are they responsible for all Spanish-language content on the platform globally?

As Mark Zuckerberg recently mentioned in testimony before Congress, our product integrity efforts are led by our VP for Product Integrity, Guy Rosen.

However, the issues around misinformation are so complex and broad, it can't fall on just one person. Developing and operationalizing policies and product updates takes a broad group of organizations and executives within the company who are involved in US Spanish-language content moderation policy and enforcement. There are diverse individuals across these teams, including Latinxs in leadership roles.

4. Phrases and symbols associated with minority communities are systematically more likely to be flagged as hate speech on social media platforms, regardless of how benign the content is. The Sikh community has noticed the targeted removal of benign posts that reference Sikh religious artifacts, political speech, and community organizing. How will Facebook create channels of recourse that will prevent the structural and systematic suppression of content from minority communities?

The conversations that happen on Facebook reflect the diversity of a community of more than two billion people communicating across countries and cultures and in dozens of languages, posting everything from text to photos and videos. We recognize how important it is for Facebook to be a place where people feel empowered to communicate. While Facebook is unequivocally opposed to hate speech and committed to removing it from our platform, we're also committed to getting better at addressing content moderation issues, including improving specific policies, our review process, and community reporting.

Our policies are enforced both automatically and by means of human review. We review and make decisions on significant numbers of pieces of content every day, and we know that we occasionally make mistakes in our initial assessments. For example, in our attempts to make sure others do not say something negative about certain groups, we sometimes incorrectly identify as hate speech content that is not hate speech. As a result, we take measures to double-check some enforcement decisions.

If someone believes that we have gotten a content moderation decision wrong, that person can generally appeal or is given the option to disagree with our decision. In most cases, we also provide users with the reason for our decision and the specific policy that was violated. In some cases, we then re-review our decisions on those individual pieces of content.

In order to request re-review of a content decision we made, users are often given the option to "Request Review" or to provide feedback by stating they "Disagree with Decision." We try to make the opportunity to request this review or give this feedback clear, either via a notification or interstitial, but we are always working to improve.

Transparency in our appeals process is important, so we now include in our Community Standards Enforcement Report how much content people appealed and how much content was restored upon appeal. Gathering and publishing those statistics keeps us accountable to the

broader community and enables us to continue improving our content moderation. For more information, see <https://transparency.facebook.com/community-standards-enforcement>.

5. Children and young adults who are members of minority communities are plagued by uniquely impactful hate speech and cyberbullying on social media platforms. While more than 15% of youth social media users experiencing cyber bullying, people of color receive disproportionately more vitriolic messages that can have substantial impact on emotional and cognitive development. How has Facebook sought to provide recourse to individuals facing online harassment at the intersection of cyberbullying and racial/ethnic hate speech?

We do not tolerate race-based attacks or harassment on our platform, and we do not permit bullying. Our Community Standards consist of over twenty policies, including policies on Hate Speech and Bullying & Harassment, developed to help make Facebook and Instagram safe. When we find content that violates our Community Standards, including those prohibiting hate speech and bullying, we remove it from Facebook. If an account continues to violate our policies, we may restrict their ability to post on Facebook or disable their profile.

Our Hate Speech policy prohibits attacks against individuals based on a number of protected characteristics, including race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity, and serious disease. Our Bullying and Harassment policies are applied to a broad range of content that attacks individuals, including content that's meant to degrade or shame. We recognize that bullying and harassment can have more of an emotional impact on minors, which is why our policies provide heightened protections for users between the ages of 13 and 18.

Facebook also encourages users to report content that may violate our policies. Our Community Operations team works hard to review those reports and takes action when content violates our policies.

6. How often do policy compliance standards change on Facebook, and who is involved in creating these policies? What is the role of international policy teams in determining global policies? How do international policies impact the policies of other nations? What safeguards does Facebook have to prevent the limitations from one country, such as India, from impacting users based in the United States?

Our Community Standards are a living set of guidelines—they must keep pace with changes happening online and in the world. The core of our policy development process is a twice-monthly, global meeting where we debate and discuss potential changes to our Community Standards. In preparation for these meetings, members of our content policy team reach out to internal and external experts, analyze data, conduct research, and study relevant scholarship to inform our policy proposals. This multi-step effort allows us to account for a range of perspectives and opinions across the globe, as well as unintended consequences and efforts to thwart our policies. When our policies are written or updated, we share those updates on our Community Standards website. More information about this process is available at <https://newsroom.fb.com/news/2019/04/insidefeed-community-standards-development-process/> and https://www.facebook.com/communitystandards/additional_information.

Our mission is to give people the power to build community and bring the world closer together—a mission that is inherently global and enhanced by a global scope. Since 2013, Facebook has been a member of the Global Network Initiative (GNI), a multi-stakeholder digital rights initiative. As part of our membership, Facebook has committed to the freedom of expression and privacy standards set out in the GNI Principles—which are in turn based on the Universal Declaration of Human Rights and the United Nations Guiding Principles on Business and Human Rights—and we are independently assessed on our compliance with these standards on a biennial basis.

In addition, our Oversight Board was created to help Facebook answer some of the most difficult questions around freedom of expression online: what to take down, what to leave up, and why. The board uses its independent judgment to support people’s right to free expression and ensure those rights are being adequately respected. The board’s decisions to uphold or reverse Facebook’s content decisions are binding, meaning Facebook will have to implement them, unless doing so could violate the law. When reviewing Facebook’s content moderation decisions, the board considers whether Facebook’s removal of the content respected international human rights standards, including on freedom of expression and other human rights. We also make it easy to identify recent updates to our Community Standards. The updates are available here: <https://www.facebook.com/communitystandards/recentupdates/>.

When governments believe that something on the internet violates local law, they may contact Facebook and ask us to remove the content. Similarly, we may receive orders to restrict content from national courts, or reports alleging illegality from non-government entities like members of the Facebook community and NGOs.

We have a robust process for reviewing reports alleging that content on Facebook or Instagram violates local law. When we receive a report, we first review it against the Facebook Community Standards or Instagram Community Guidelines. If we determine that the content violates our policies, we remove it. If content does not violate our policies, in line with our commitments as a member of the Global Network Initiative and our Corporate Human Rights Policy, we conduct a careful legal review to confirm whether the report is valid, as well as human rights due diligence.

In cases where we believe that reports are not legally valid, are overly broad, or are inconsistent with international human rights standards, we may request clarification or take no action. In all cases, we consider the impact our decisions will have on the availability of other speech via our products.

Where we do act against organic content on the basis of local law rather than our Community Standards, we restrict access to the content only in the jurisdiction where it is alleged to be unlawful and do not impose any other penalties or feature restrictions. We provide notice to people when we restrict something they posted based on a report of an alleged violation of local law, and we also tell people when they try view something that is restricted in their country on the basis of a legal request.

7. Can you explain what exactly you have done to ensure Facebook’s tools — algorithms, recommendation engines, and targeting tools — are not amplifying

conspiracy theories and disinformation, connecting people to dangerous content, or recommending hate groups or purveyors of disinformation and conspiracy theories to people? For example, can you provide detailed answers on some of the Capitol riot suspects' use history, to include the following:

- a. Which Facebook groups were they members of? Did Facebook recommend those groups, or did the individuals search for the specific groups on their own? What ads were targeted at them based on either the data you gathered or interests you inferred about them? Were they connected to any known conspiracy theorists, QAnon believers, or other known January 6th rioters due to Facebook's recommendations?**

We are unable to share specific information about the Capitol riot suspects at this time. We can tell you we are working closely with law enforcement as they investigate events surrounding January 6. We dedicate significant resources to addressing law enforcement concerns, and we carefully review, validate, and respond to the legal requests we receive from them as soon as possible. This includes prioritizing requests related to emergency situations, including those we have received from federal and state law enforcement agencies related to the Capitol attack. And when we see a credible threat on our platform, we reach out to law enforcement proactively.

We are also proud of the work that we have done to keep hate and violence off of our platform. While our enforcement efforts are not perfect—and there is always more work to be done—we have taken strong steps to keep our community safe.

Our Dangerous Organizations and Individuals policy prohibits content calling for or advocating violence, and we ban organizations and individuals that proclaim a violent mission. We believe this policy has long been the broadest and most aggressive in the industry. In August 2020, we expanded this policy further to address militarized social movements and violence-inducing conspiracy networks, such as QAnon. And in October of last year, we began removing any Facebook Pages, Groups, or Instagram accounts representing QAnon, regardless of whether they contained violent content. To date, we've banned over 250 white supremacist groups and have been enforcing our rules that prohibit QAnon and militia groups from organizing on our platform. We have also continued to enforce our ban on hate groups, including the Proud Boys and many others.

We also recognize the importance of keeping violent and hateful content out of Groups and have taken strong and decisive actions to work towards that goal. We remove Facebook Groups that represent QAnon, even if the Groups contain no violent content. And we do not allow militarized social movements, such as militias or groups that support and organize violent acts amid protests, to have a presence on our platform. Last year, we also temporarily stopped recommending US civic or political Groups to our users. Earlier this year, we announced that policy would be kept in place and expanded globally.

- 8. After witnessing the tragic events in Kenosha, what will Facebook do to ensure someone cannot create a Facebook page calling for armed individuals to organize a militia? How will Facebook enforce this policy?**

The events that took place in Kenosha last year were a national tragedy, and we work hard to prevent our platforms from being used in any way that might lead to such an event.

Over the last several years, we've continued to update and refine our Violence and Incitement policy, including in August 2020, when we developed a framework whereby we can identify certain locations that are more at risk for violence or intimidation by the threat of violence, the same as we identify schools, polling places, and houses of worship, and remove more implicit calls and statements to bring weapons to those locations.

We also updated our policies to address militarized social movements and violence-inducing conspiracy networks, such as QAnon. To date, we have banned over 250 white supremacist groups and 890 militarized social movements from Facebook and Instagram, and we have removed tens of thousands of Pages and Groups related to militarized social movements and QAnon. We have also continued to enforce our ban on hate groups, including the Proud Boys and many others.

9. Facebook is clearly not enforcing its prohibition on microtargeting ads towards persons with disabilities. As a result, not only are children being harmed, but charlatans are empowered to monetize their fake cures. What commitments can you make to ensuring this ad policy is enforced?

We do not have targeting options based on medical condition or disabilities, and we prohibit businesses from sending us sensitive health information. Over the years, we've been removing targeting options to help prevent misuse. While these options have been used in legitimate ways to reach people interested in a certain product or topic, we think minimizing the risk of abuse is more important. For example, we've removed some interests that directly describe a medical condition, but we have kept some health-related interests that we believe help promote awareness or support around health issues.

We are committed to making our platforms safe. Under our Misinformation and Harm policy, we remove misinformation that contributes to the risk of imminent violence or physical harm. Our prohibitions against false and deceptive claims in advertisements encompass prohibitions on misleading health claims such as promotions of miracle cures.

10. Facebook recently announced that it would start sending retroactive corrections to users who interacted with misinformation about COVID-19 that was subsequently removed from the platform. How will you ensure enforcement of this policy and measure your compliance? Will Facebook provide users with transparency about when they have seen false content by offering retroactive corrections to posts fact-checked by your partners?

As your question notes, we send messages to people who have interacted with misinformation about COVID-19 on Facebook that we've since removed. In April 2020, we started showing these messages in News Feed to people who liked, commented on, or reacted to posts with misinformation that we removed for violating our policy. Since then, we've done research to better understand what's most helpful for people, and we've redesigned these as more personalized notifications to more clearly connect people with credible and accurate information about COVID-19.

As a result, people now receive a notification that says we've removed a post they've interacted with for violating our policy against misinformation about COVID-19 that leads to imminent physical harm. Once they click on the notification, they will see a thumbnail of the post and more information about where they saw it and how they engaged with it. They will also see why it was false and why we removed it (e.g., because the post included the false claim that COVID-19 doesn't exist). People will then be able to see more facts about COVID-19 in our Coronavirus Information Center and take other actions, such as unfollowing the Page or Group that shared this content.

11. In October 2020, Facebook announced it was removing militarized social movements such as QAnon from its network. Before this prohibition, members of QAnon were free to organize, network, and recruit on Facebook, and Facebook sold advertising as they did so.

a. How much revenue did Facebook generate from the QAnon movement before this prohibition?

We commit significant resources to enforcing our Community Standards and prioritizing safety on our platforms. Our Dangerous Organizations and Individuals policy prohibits content calling for or advocating violence, and we ban organizations and individuals that proclaim a violent mission. We believe this policy has long been the broadest and most aggressive in the industry. As your question notes, in August 2020, we expanded this policy further to address militarized social movements and violence-inducing conspiracy networks, such as QAnon. In September, we prohibited anyone on our platform from running ads that praise, support, or represent militarized social movements and QAnon. And in October, we began removing any Facebook Pages, Groups, or Instagram accounts representing QAnon, regardless of whether they contained violent content.

We remove language that incites or facilitates violence, and we ban groups that proclaim a hateful and violent mission from having a presence on our apps. We also remove content that represents, praises, or supports those groups. To date, we have banned over 250 white supremacist groups and 890 militarized social movements from Facebook and Instagram, and we have removed tens of thousands of Pages and Groups related to militarized social movements and QAnon. We have also continued to enforce our ban on hate groups, including the Proud Boys and many others.

b. In your 2020 annual report, you warned investors that user loss was a significant investment risk. "A number of other social networking companies that achieved early popularity," you wrote, "have since seen their active user bases or levels of engagement decline, in some cases precipitously." Are you concerned about the loss of users when making content moderation decisions?

c. Did you do any financial modeling or revenue projections when making content moderation decisions around militarized social movements? Can you share those with the committee? Are you tracking user loss in the wake of these decisions? Can you share that internal analysis with the committee?

Billions of people use Facebook and Instagram because they have good experiences; they don't want to see hateful and violent content, our advertisers don't want to see it, and we don't want to see it. There is no incentive for us to do anything but remove it. We are committed to doing our part to keep people safe on our services and protect free expression, and we work hard to set and enforce policies that meet those goals. We will continue to invest extraordinary resources into content moderation, enforcement, and transparency.

Our Dangerous Organizations and Individuals policy prohibits content calling for or advocating violence, and we ban organizations and individuals that proclaim a violent mission. We believe this policy has long been the broadest and most aggressive in the industry, and it predates our October 2020 announcement regarding militarized social movements. For instance, Facebook removed the Proud Boys from the platform in 2018. In August 2020, we expanded this policy further to address militarized social movements and violence-inducing conspiracy networks, such as QAnon. We remove language that incites or facilitates violence, and we ban groups that proclaim a hateful and violent mission from having a presence on our apps. We also remove content that represents, praises, or supports those groups.

To date, we have banned over 250 white supremacist groups and 890 militarized social movements, and we have been enforcing our rules that prohibit QAnon and militia groups from organizing on our platform. We have also continued to enforce our ban on hate groups, including the Proud Boys and many others.

We also recognize the importance of keeping violent and hateful content out of Groups and have taken strong and decisive actions to work towards that goal. We remove Facebook Groups that represent QAnon, even if the Groups contain no violent content. And we do not allow militarized social movements, such as militias or groups that support and organize violent acts amid protests, to have a presence on our platform. Last year, we also temporarily stopped recommending US civic or political Groups to our users. Earlier this year, we announced that policy would be kept in place and expanded globally.

12. Did you do any financial modeling or revenue projections when making content moderation decisions around Donald Trump's presence on Facebook?

a. Can you share those with the committee? Are you tracking user loss in the wake of these decisions? Can you share that internal analysis with the committee?

Please see the response to your Question 11.

We are committed to keeping people safe on our services and to protecting free expression, and we work hard to set and enforce policies that meet those goals. We will continue to invest extraordinary resources into content moderation, enforcement, and transparency.

Our decision to place a feature block on former President Trump's account was made consistent with our Community Standards, which inform how we moderate content. In this extraordinary case, we determined that former President Trump's posts, on balance, contributed to, rather than diminished, the risk of ongoing violence.

- 13. Research from InfluenceMap has shown that Facebook advertising has been used by organizations to pay to spread and target climate change disinformation. InfluenceMap has also recently noted cases where Facebook has failed to classify ads containing climate disinformation as political or issue ads during, and after, its recent political ad ban. Will Facebook commit to eliminating all climate change disinformation spread through the Facebook advertising?**

We recognize the urgency of climate change and are committed to helping tackle this global challenge. Our independent third-party fact-checking partners review and rate climate misinformation. In fact, we have a fact-checking partner specifically dedicated to reviewing science content, and many of the other partners in our network of over 80 global fact-checking partner organizations rate this content as well. As with all claims debunked by our partners, we reduce the distribution of posts rated False, Altered, or Partly False in News Feed and apply a label on top of them on Facebook and Instagram, so people understand that the content has been rated and what the ratings mean.

We reject ads that include content that has been debunked on our platform by our fact-checking partners. We are the only company to partner with 80 fact-checking organizations, using AI to scale their fact-checks to millions of duplicate posts.

Additionally, ads that concern issues related to climate change, like all ads on our platforms, are subject to Facebook's ad review system, which relies primarily on automated review to check ads against our Community Standards and Advertising Policies. These policies are posted publicly online. We use teams of reviewers to improve and train our automated systems, and in some cases, to review specific ads, and we have added several thousand reviewers over the past few years as we have expanded our coverage of ad review.

Please see the response to your Question 14 below for more information about our fact-checking efforts specifically with regard to climate change.

- 14. As recently as February, misinformation about wind turbines causing the power outages in Texas spread seemingly unchecked on Facebook. Will you commit to monitoring climate change disinformation, and releasing regular and transparent reports to the public?**

As described in the response to your prior question, our independent third-party fact-checking partners can review and rate climate misinformation. Additionally, our Climate Science Information Center on Facebook provides persistent access to global, regional, and local authoritative information about climate change and its effects, similar to what we've done with COVID-19 and elections. The Center connects people to factual, timely, regionally relevant information wherever it makes sense in the Facebook experience. This is the first step in a dedicated effort by our Discovery and Sustainability teams. The Center features resources from the world's leading climate organizations and clear steps people can take to combat climate change. We're working with the Intergovernmental Panel on Climate Change (IPCC) and their global network of climate science contributors to include facts, figures, and data. Contributors include the UN Environment Programme (UNEP), the National Oceanic and Atmospheric

Administration (NOAA), and the World Meteorological Organization (WMO). We also include posts from relevant sources to highlight climate science news.

We recently announced improvements to the Climate Science Information Center, expanding its reach and providing additional science-based news to our users. We added a section that features facts that debunk common climate myths. To debunk the myths with current and specific facts, we've brought in climate communication experts from George Mason University, the Yale Program on Climate Change Communication, and the University of Cambridge.

15. It's not clear what counts as a strike when it comes to climate disinformation. Will Facebook adopt clear rules around strikes for repeat offenders of climate disinformation?

Please see the response to your previous question.

16. In the weeks immediately before the 2020 election, Facebook reportedly sought to improve the quality of News Feed content by emphasizing in its ranking algorithm News Ecosystem Quality (NEQ) scores—internal rankings Facebook gives publishers based on its journalistic assessment.¹ Facebook then deemphasized NEQ in the News Feed ranking algorithm following the January 6 insurrection, but has since indicated the temporary change has been rolled back. However, recent analysis shows this and other “emergency” measures taken around the election were quite effective in reducing misinformation spread.

- a. Please explain why you waited so long to implement and then so quickly removed these measures. Please provide any general criteria Facebook uses in making these decisions.**
- b. Why do you view the level of NEQ emphasis in the News Feed algorithm and the other actions Facebook took surrounding the election/insurrection as temporary, emergency actions?**

Facebook worked hard to do our part in helping to protect the integrity of the 2020 election, and we're proud of the work we've done to support our democracy. We followed the policies and processes that we laid out in advance to protect the democratic process in both the pre- and post-election periods, and we worked hard to apply those policies fairly and consistently. Securing the integrity of elections is an ongoing challenge for platforms, and we are committed to continuing to improve our systems.

At Facebook, we often have to balance competing equities. Sometimes the right thing to do from a safety or security perspective isn't the best for privacy or free expression. Making these trade-offs is not straightforward, so we have to choose what we believe is best for our community and for the world.

¹ Facebook Struggles to Balance Civility and Growth, *The New York Times* (Nov. 24, 2020), www.nytimes.com/2020/11/24/technology/facebook-election-misinformation.html.

The measure referenced in your question was a temporary change we made to help limit the spread of inaccurate claims about the election, and it was part of our elections integrity efforts. We said at the time that these measures were intended to be temporary. In fact, while political content in an election season represents only around 6% of what people see, we have seen even that small percentage decline now that the election is over.

News stories in general make up a small portion of what people see on Facebook overall, and political news makes up a smaller fraction of that news content. We're still ensuring that people see authoritative and informative news on Facebook, especially during major news cycles and around important global topics like elections, COVID-19, and climate change.

- 17. Along with law enforcement agencies, Congress is conducting multiple lines of inquiry into January 6th and there may indeed be a National Commission that will have the role of the tech platforms in its remit. Have you taken proactive steps to preserve evidence that may be relevant to the election disinformation campaign that resulted in the January 6th siege on the Capitol, and to preserve all accounts, Groups, and exchanges of information on your sites that may be associated with parties that participated in it?**

We are committed to continuing to work with law enforcement and Congress as they investigate the attacks on the Capitol. As part of this work, we are actively responding to valid legal process from federal and local law enforcement authorities, as well as proactively sharing information when consistent with the law and our terms of service. And, in connection with these ongoing efforts, we are preserving the content related to valid law enforcement requests and the information we proactively share.

- 18. Mr. Zuckerberg, you and other Facebook executives have routinely testified to lawmakers and regulators that Facebook's AI finds and removes as much as 99% of some forms of objectionable content, such as terrorist propaganda, human trafficking content and, more recently, child sex exploitation content. It is normally understood to mean that Facebook AI and moderators remove 99% of overall content. But can you define clearly that you mean to say your AI removes 99% of what you remove, rather than the total amount of such content? 8 Does Facebook have evidence about its overall rate of removal of terrorist content, human trafficking content, and child sexual abuse material (CSAM) that it can provide to this Committee? Studies by the Alliance to Counter Crime Online indicate you are removing only about 25-30%. Can you explain the discrepancy?**

We publish our Community Standards Enforcement Report (<https://transparency.facebook.com/community-standards-enforcement>) on a quarterly basis to more effectively track our progress and demonstrate our continued commitment to making Facebook and Instagram safe and inclusive. The Report includes a variety of metrics, including the number of pieces of violating content actioned in each category and the percentage of that content that we found and flagged before users reported it.

We have also been working to add a prevalence metric, which is currently available for many of the categories in the Report. Prevalence considers all the views of content on Facebook or Instagram and measures the estimated percentage of those views that were of violating

content. This metric assumes that the impact caused by violating content is proportional to the number of times that content is viewed.

Another way to think of prevalence is how many views of violating content we didn't prevent—either because we haven't caught the violations early enough or we missed them altogether. We hold ourselves accountable to this number. You can learn more about prevalence in our blog post “Measuring Prevalence of Violating Content on Facebook” (<https://newsroom.fb.com/news/2019/05/measuring-prevalence/>).

Prevalence of violating content is estimated using samples of content views from across Facebook or Instagram. We calculate it as: the estimated number of views that showed violating content, divided by the estimated number of total content views on Facebook or Instagram. If the prevalence of adult nudity and sexual activity was 0.18% to 0.20%, that would mean of every 10,000 content views, 18 to 20 on average were of content that violated our standards for adult nudity and sexual activity.

Some types of violations occur very infrequently on our services. The likelihood that people view violating content is very low, and we remove much of that content before people see it. As a result, many times we do not find enough violating samples to precisely estimate prevalence. In these cases, we can estimate an upper limit of how often someone would see content that violates these policies. For example, if the upper limit for terrorist propaganda was 0.04%, that means that out of every 10,000 views on Facebook or Instagram in that time period, we estimate that no more than 4 of those views contained content that violated our terrorist propaganda policy.

In the fourth quarter of 2020, the upper limit was 0.07% for violations of our policy for terrorism on Facebook. This means that out of every 10,000 views of content on Facebook, we estimate no more than seven of those views contained content that violated the policy. In the fourth quarter of 2020, the upper limit was 0.05% for violations of our policy for child nudity and sexual exploitation of children on Facebook. This means that out of every 10,000 views of content on Facebook, we estimate no more than five of those views contained content that violated the policy.

It's important to note that when the prevalence of a violation type is so low that we can only provide upper limits, this limit may change by a few hundredths of a percentage point between reporting periods. However, changes this small may not be statistically significant; in such cases, these small changes do not indicate an actual difference in the prevalence of this violating content on the service.

We estimate how often content is seen rather than the amount of content posted because we want to determine how much that content affected people on Facebook or Instagram. A piece of violating content could be published once but seen a thousand times, a million times, or not at all. Measuring views of violating content rather than the amount of violating content published better reflects the impact on the community. A small prevalence number can still correspond to a large amount of impact on our services, due to the large number of overall views of content on our services.

We estimate prevalence by sampling content views on Facebook or Instagram. To do this, we manually review samples of views and the content shown in them. Then we label the samples as violating or not violating according to our policies. Using the portion of these samples that were of violating content, we estimate the percentage of all views that were of violating content. Note that we do not sample from every part of Facebook or Instagram for every violation type.

For violation types that are viewed very infrequently, sampling requires a very large number of content samples to estimate a precise prevalence measure. For these types of violations, rather than use stratified sampling, we do random sampling. In these cases, we can only estimate the upper limit—meaning that we are confident that the prevalence of violating views is below that limit, but we cannot precisely say how far below.

19. Former Facebook policy employees came forward to say that “Mark personally didn’t like the punishment, so he changed the rules,” when it came to banning Alex Jones and other extremists like the Oath Keepers. Mr. Zuckerberg, what role do you play in the moderation of misinformation and deciding what harmful content qualifies for removal?

First, Alex Jones was banned from Facebook in 2019. And in fall 2020, the Oath Keepers were designated a militarized social movement under our Dangerous Organizations and Individuals policy and have also been banned from Facebook.

Big decisions at Facebook are made with input from people across different teams who have different perspectives and expertise in different areas, but ultimately Mark Zuckerberg is responsible as the leader of the company for decisions that are made. And we’ve always banned individuals or organizations that promote or engage in violence and hate, regardless of ideology.

20. A recent *BuzzFeed News* report found that following the November 2020 election at least six Facebook employees resigned and posted farewells criticizing Facebook leadership on misinformation and hate speech.² Four of these employees explicitly cited the company’s organizational structure requiring the content policy team that creates and enforces platform rules to report directly to the head of the public policy division that deals with governments and regulators worldwide as a key failure in content moderation.

a. Please confirm whether this structure remains in place as of April 8, 2021. If not, please describe in detail what changes were made and why.

Our content policy team reports to our head of public policy.

b. Is it Facebook’s policy to seek political balance in the outcomes of its content moderation decisions?

² “Mark Changed The Rules”: How Facebook Went Easy On Alex Jones And Other Right-Wing Figures, *BuzzFeed News* (Feb. 21, 2021), www.buzzfeednews.com/article/ryanmac/mark-zuckerberg-joel-kaplan-facebook-alex-jones.

Facebook is a platform for ideas across the political spectrum. We are committed to ensuring there is no bias in the work we do. Suppressing content on the basis of political viewpoint or preventing people from seeing what matters most to them is directly contrary to Facebook’s mission and our business objectives.

- c. Please explain the reasoning behind this structure for content moderation, which is distinct among industry leaders. Did Facebook consider the potential for misaligned incentives between content moderation rules and public policy initiatives? Are changes under consideration?**

Both our Content Policy team and the broader Public Policy team are guided by the same core principles, including the values of authenticity, safety, privacy, and dignity that are the foundation of our Community Standards.

- d. Do you agree that this structure contributed to Facebook’s failure to respond in a timely and effective manner to election disinformation proliferating on the platform before the 2020 election and contributing to the January 6 insurrection? If so, how?**

The Capitol attack was a horrific assault on our values and our democracy, and Facebook is committed to assisting law enforcement in bringing the insurrectionists to justice. We are also committed to keeping people safe on our services and to protecting free expression, and we work hard to set and enforce policies that meet those goals. We will continue to invest extraordinary resources into content moderation, enforcement, and transparency.

Facebook worked hard to do our part in protecting the integrity of the 2020 election, and we’re proud of the work we’ve done to support our democracy. We followed the policies and processes that we laid out in advance to protect the democratic process in both the pre- and post-election periods, and we worked hard to apply those policies fairly and consistently. Securing the integrity of elections is an ongoing challenge for platforms, and we are committed to continuing to improve our systems.

We are also proud of the work that we have done to keep hate and violence off of our platform. While our enforcement efforts are not perfect—and there is always more work to be done—we have taken strong steps to keep our community safe.

Our Dangerous Organizations and Individuals policy prohibits content calling for or advocating violence, and we ban organizations and individuals that proclaim a violent mission. We believe this policy has long been the broadest and most aggressive in the industry. In August 2020, we expanded this policy further to address militarized social movements and violence-inducing conspiracy networks, such as QAnon. We remove language that incites or facilitates violence, and we ban groups that proclaim a hateful and violent mission from having a presence on our apps. We also remove content that represents, praises, or supports those groups.

To date, we have banned over 250 white supremacist groups and 890 militarized social movements from Facebook and Instagram, and we have removed tens of thousands of Pages and Groups related to militarized social movements and QAnon. We have also continued to enforce our ban on hate groups, including the Proud Boys and many others.

We also recognize the importance of keeping violent and hateful content out of Groups and have taken strong and decisive actions to work towards that goal. We remove Facebook Groups that represent QAnon, even if the Groups contain no violent content. And we do not allow militarized social movements, such as militias or groups that support and organize violent acts amid protests, to have a presence on our platform. Last year, we also temporarily stopped recommending US civic or political Groups to our users. Earlier this year, we announced that policy would be kept in place and expanded globally.

21. 2016 research from Facebook showed that 64% of people who joined FB groups promoting extremist content did so at the prompting of Facebook’s recommendation tools. Facebook reportedly changed its policies. You were previously asked in a Senate hearing whether you had seen a reduction in your platform’s facilitation of extremist group recruitment since those policies were changed, to which you responded, “Senator, I’m not familiar with that specific study.” Are you now familiar with that study, and what’s your response now — did you see a reduction in your platform’s facilitation of extremist group recruitment since those policies were changed?

As we previously responded to Senator Peters, when asked the question referenced above in the prior hearing, Mr. Zuckerberg did not immediately recall the study Senator Peters was referencing. The document referenced related to a specific issue identified in Germany ahead of the 2017 German elections, which we addressed at the time.

Any suggestion that we have buried research on this topic or haven’t acted on these important issues is inaccurate. For example, in 2018, we responded to feedback from our community that public content—posts from businesses, brands, and media—was crowding out the personal moments that lead us to connect more with one another. As a result, we moved from focusing solely on helping users find relevant content to helping them have more meaningful social interactions. This meant that users began seeing more content from their friends, family, and Groups. We also reduce the distribution of some problematic types of content, including content that users may find spammy or low-quality, such as clickbait headlines and links to low-quality webpages like ad farms.

We also fund research on misinformation and polarization to better understand the impact of our products; for example, in February 2020, we announced an additional \$2 million in funding for independent research on this topic.

We are proud of the work we have done to make Facebook an unwelcome place for those committed to acts of violence. In fact, our Dangerous Organizations and Individuals policy has long been the broadest and most aggressive in the industry. And in August 2020, we expanded that policy to address militarized social movements and violence-inducing conspiracy networks, such as QAnon. The purpose of this policy is to prevent offline harm that may be related to content on Facebook, and so in the course of that work we contact law enforcement if we see imminent credible threats on the platform. We remove language that incites or facilitates serious violence. We also ban groups that proclaim a hateful and violent mission from having a presence on our apps, and we remove content that represents, praises, or supports them.

Moving fast to find and remove dangerous organizations, including terrorist and hate groups, takes significant investment in both people and technology. At Facebook, we have tripled the size of our teams working in safety and security since 2016 to over 35,000 people—including teams that review reports of hate speech and content that praises, supports, or represents hate groups. We also have several hundred people who exclusively or primarily focus on countering dangerous organizations as their core responsibility. This group includes former academics who are experts on counterterrorism, former prosecutors, former law enforcement agents, investigators, analysts, and engineers.

Four years ago, we developed a playbook and a series of automated techniques to detect content related to terrorist organizations such as ISIS, al Qaeda, and their affiliates. We've since expanded these techniques to detect and remove content related to other terrorist and hate groups. We're now able to detect text embedded in images and videos in order to understand its full context, and we've built media-matching technology to find content that's identical or near-identical to photos, videos, text, and even audio that we've already removed.

When we started detecting hate organizations, we focused on groups that posed the greatest threat of violence at that time, and we've now expanded to detect more groups tied to different hate-based and violent extremist ideologies and using different languages. In addition to building new tools, we've also adapted strategies from our counterterrorism work, such as leveraging off-platform signals to identify dangerous content on Facebook and implementing procedures to audit the accuracy of our AI's decisions over time.

We understand, however, that simply working to keep violence off Facebook is not an adequate solution to the problem of online content tied to violent extremism, particularly because bad actors can leverage a variety of platforms and operate offline as well. While we work 24/7 to identify, review, and remove violent extremist content, our efforts do not stop there. We believe our partnerships with other companies, civil society, researchers, and governments are crucial to combating this threat. For example, our P2P Global Digital Challenge, which engages university students around the world in competitions to create social media campaigns and offline strategies to challenge hateful and extremist narratives, has launched over 600 counterspeech campaigns from students in 75 countries, engaged over 6,500 students, and reached over 200 million people. We have also developed the Redirect Initiative to connect people searching for violent extremist material with offline organizations dedicated to helping people disconnect from extremist groups. The program is active now in four countries, including the US, where we have partnered with Life After Hate, an organization founded by former violent extremists, to help people disconnect from white supremacist groups.

In addition, in 2017 we helped establish the Global Internet Forum to Counter Terrorism (GIFCT) with others in the industry, with the objective of disrupting terrorist abuse on our platforms. Since then, the consortium has grown and collaborates closely on critical initiatives focused on tech innovation, knowledge-sharing, and research. In 2019, we reached our goal of collectively contributing more than 200,000 hashes, or unique digital fingerprints, of known terrorist content into our shared database, enabling each of us to quickly identify and take action on potential terrorist content on our respective platforms.

22. Little to no information has been released by Facebook regarding why or how #Sikh and #Sikhism were blocked multiple times throughout 2020. What procedures does your platform follow to remove or block hashtags?

a. Were automated systems responsible for the initial takedowns of these hashtags or was human decision-making involved in these processes?

We investigated this issue and found that these hashtags were mistakenly blocked on March 7, 2020 following a report that was inaccurately reviewed by our teams. We became aware that the hashtags were blocked on June 3, 2020, following feedback we received from the community, and we quickly moved to unblock them. Our processes fell down here, and we're sorry. We understand that this is an incredibly important, painful time for the Sikh community. We designed hashtags to allow people to come together and share with one another. It's never our intention to silence the voices of this community, and we are taking the necessary steps so this doesn't happen again.

23. Your testimony cited Facebook's success at removing COVID-19 and vaccine-related disinformation. At the same time, white supremacist content, racist, and other hate content is running rampant on the platform. Is there anything inherently more difficult about policing white supremacist, racist, and other hate content compared to health misinformation and disinformation? Why?

When we find hateful posts on Facebook and Instagram, we take a zero-tolerance approach and remove them. In fact, we invest billions of dollars each year in people and technology to keep our platform safe. We have tripled—to more than 35,000—the people working on safety and security, and we are a pioneer in artificial intelligence technology to remove hateful content at scale.

We similarly have no tolerance for violence, and we are proud of the work we have done to make Facebook an unwelcome place for those committed to acts of violence. In fact, our Dangerous Organizations and Individuals policy has long been the broadest and most aggressive in the industry. And in August 2020, we expanded that policy to address militarized social movements and violence-inducing conspiracy networks, such as QAnon. We remove language that incites or facilitates serious violence. We also ban groups that proclaim a hateful and violent mission from having a presence on our apps, and we remove content that represents, praises, or supports them. We also have several hundred people who exclusively or primarily focus on countering dangerous organizations as their core responsibility. This group includes former academics who are experts on counterterrorism, former prosecutors, former law enforcement agents, investigators, analysts, and engineers.

Zero tolerance does not mean zero incidences, but we're making real progress. A recent European Commission report found that Facebook assessed 95.7 percent of hate speech reports in less than 24 hours. In February, we reported that we find more than 97 percent of the hate speech we remove before someone reports it—up from 24 percent three years earlier. We took action against 26.9 million pieces of hate speech content in the last quarter of 2020—up from 22.1 million in the previous quarter. More than 99 percent of the terrorist content we remove is taken down before anyone reports it to us, along with more than 98 percent of the organized hate content.

We also measure prevalence of violating content to gauge how we're performing against our goal to minimize the impact caused by violations of our policies on people using our services. Prevalence considers all the views of content on Facebook or Instagram and measures the estimated percentage of those views that were of violating content. In the fourth quarter of 2020, we estimate that between 0.07% and 0.08% of views on Facebook were of content that violated our Hate Speech policy, meaning that hate speech content was viewed 7 or 8 times for every 10,000 views of content on Facebook. This was a decrease from the previous quarter, which was due to ranking changes we made to personalize content for users and reduce problematic content in News Feed. We will be releasing updated metrics for the first quarter of 2021 later this month in our next Community Standards Enforcement Report, which will be available at <https://transparency.facebook.com/community-standards-enforcement>.

24. The personal data of over 533 million Facebook users was recently posted in a hacking forum and made available to the entire world. This information included users' Facebook identifiers, phone numbers, full names, locations, birth dates, email addresses, biographical information, and other personal information. According to a Reuters report, Facebook has not and does not intend to notify any of these users whose personal information was stolen.

a. Facebook claims that the data involved was scraped from the platform before September 2019. What evidence does Facebook have that the malicious actors scraped the data and not through hacking Facebook's system?

We have been working with external security researchers to analyze the scraped data set that is the subject of this reporting (the "Scraped Data Set"). Based on that analysis, we concluded that the Scraped Data Set was compiled using a scraping method known as phone number enumeration. This scraping technique involves inputting or uploading a large set of phone numbers to see which ones return information that is already publicly available using features designed for, and valued by, people who use our services. These features work according to user privacy permissions; that is, they only returned contacts with matching phone numbers if those contacts' privacy settings permitted others to search for them by phone number.

b. Facebook also claimed the particular data-scraping vulnerability allowing for this data breach was repaired in 2019. What evidence does Facebook have that none of this data was obtained after September 2019?

Based on the security researchers' ongoing analysis, Facebook believes that the Scraped Data Set was compiled no later than September 2019, through phone number enumeration scraping using contact discovery features on Facebook's platforms. These features were designed to provide an easy way for users to find people they know on Facebook platforms, including by searching for someone via the person's phone number, or uploading contacts from a device to see which ones are on Facebook. These features have at all relevant times been limited by user privacy permissions; that is, they have only returned contacts with matching phone numbers if those contacts' privacy settings permitted others to search for them by phone number.

Before 2019, Facebook imposed measures to deter and limit scraping on these features, including through rate limits. These efforts were designed to accommodate the full range of

legitimate usage of these features, while making it difficult for malicious actors to exploit them. However, Facebook believes that the scrapers responsible for the Scraped Data Set evaded these rate limits by using bots to emulate large numbers of people using contact discovery features. In this way, malicious actors could conduct a large number of phone number lookups in order to find matching users whose phone number search settings were set to “everyone.” While the total usage by each bot would stay below applicable rate limits, the lookups could nonetheless be run at scale through the parallel use of the numerous bots.

Over the course of 2018 and 2019, as scrapers modified their tactics, Facebook adapted to these changes in behavior by modifying the contact discovery features. These changes made the search and contact importer features less useful for Facebook users, but Facebook believes that the changes mitigated the scraping methods that were likely used to assemble the Scraped Data Set. The last of these changes was made in September 2019, which also appears to be the end of the date range of the records in the Scraped Data Set.

- c. Can user data be scraped only from users who have their audience or privacy settings set to public? Can user data be scraped from users that have set privacy and audience settings to “friends only,” “friends except,” “specific friends,” or “only me”?**

The Scraped Data Set contains public information from Facebook users’ profiles combined with phone numbers. The phone numbers are believed to have been supplied by malicious actors as part of the phone number enumeration scraping they conducted to identify the corresponding users. The data scraped from the users’ profiles consists of limited categories of profile information that were publicly accessible on Facebook.

The contact discovery features at issue here were all relevant times limited by user privacy permissions: that is, the contact discovery features only returned user information to matching phone numbers if those users’ privacy settings permitted “everyone” to look them up by phone number, and they only provided user information in response to a phone number query consistent with that user’s audience settings. This means that scrapers with no relationship to a user would only have been able to obtain information from users whose privacy settings allowed “everyone” to look them up by phone number, and then they only returned information that the users shared publicly on Facebook.

- d. Does Facebook contend it is not obligated under any state breach notification law to provide notice to any of these users whose data was improperly scraped from your platform and posted publicly without user authorization?**

The Scraped Data Set was not obtained through a security breach. As explained, it was obtained through scraping publicly accessible information. Consistent with industry practice, Facebook does not consider such scraping activity to constitute a reportable security breach.

- e. Facebook promised to bolster its protections against mass data-scraping after the Cambridge Analytica scandal. The 2019 amendments to Facebook’s consent order with the FTC defined a “Covered Incident” as “any instance in which Respondent has verified or otherwise confirmed that**

the Covered Information of 500 or more Users was or was likely to have been accessed, collected, used, or shared by a Covered Third Party in violation of Respondent’s Platform Terms.” Did this scraping incident reported by Business Insider on April 3 violate Facebook’s Terms?

Section 3.2.3 of Facebook’s Terms of Service prohibits collecting data from Facebook’s products using automated means without Facebook’s permission. Because the Scraped Data Set was obtained through the unauthorized use of automated means of collecting data from our products, the incident constitutes a violation of our Terms.

25. Last month Facebook ended its ban on political advertising that followed the 2020 election. Facebook, however, did not revise its policies on political ads or increasing enforcement against violations of existing policies.

- a. Please explain what, if any, changes Facebook has made to its policies related to political advertising since the 2020 election.**
- b. What efforts are Facebook currently taking or planning to take to determine what changes to its political advertising policy are necessary to combat the spread of harmful misinformation and disinformation?**

We resumed political, electoral, and social issue ads in the United States on March 4, 2021. At that time, we also expanded our ad control for people to see fewer ads about social issues, elections, or politics with “paid for by” disclaimers in 90+ countries and territories globally. We put a temporary ban on such ads in place after the November 2020 election to avoid confusion or abuse following Election Day. Unlike other platforms, we require authorization and transparency not just for political and electoral ads, but also for social issue ads, and our systems do not distinguish between these categories. We’ve heard a lot of feedback about this and learned more about political and electoral ads during last year’s election cycle. As a result, we plan to use the coming months to take a closer look at how these ads work on our service to see where further changes may be merited.

26. Even before COVID-19 and as recently as January 2020, anti-vaccination groups and profiteers ran ads on Facebook leading to resurgences in diseases like measles that our country had all-but-eradicated.

- a. What were Facebook’s estimated profits from anti-vaccination ad content pre-COVID-19?**
- b. How have these numbers changed during the pandemic?**
- c. How many clicks for anti-vaccination conspiracy advertisements has Facebook generated?**
- d. How much revenue did Facebook generate from allowing advertisers to target people deemed to be interested in “pseudoscience”?**

Ads that include misinformation about vaccines have been prohibited on Facebook since 2019, irrespective of who is running the ad.

In December, we began removing from our platform false claims about COVID-19 vaccines that could lead to imminent harm, including false claims about the safety, efficacy, ingredients, or side effects of the vaccines. On February 8, we expanded the list of false claims we will remove to include additional debunked claims about COVID-19 and vaccines following consultations with leading health organizations, including the World Health Organization (WHO).

The Honorable Bobby L. Rush

- 1. Instantly notify and provide corrections to every user who has been exposed to content determined by Facebook’s fact-checking partners to be false and misleading? If so, what is the expected timeline for this to take place? If not, why not?**

We’re committed to fighting the spread of false information on Facebook. We use both technology and human review to remove fake accounts, promote information literacy, and disrupt the financial incentives of spammers. In many countries, we also work with independent fact-checkers who are certified through the nonpartisan International Fact-Checking Network to help identify and review false information.³

When fact-checkers write articles with more information about a story, you’ll see a label on the post and have an option to click to see why. You’ll also receive a notification if you try to share a story or have shared one in the past that’s been rated false by fact-checkers. Page Admins will also be notified if they share stories rated false.⁴ For misinformation that violates our Community Standards, such as false claims about COVID-19 and vaccines that lead to imminent physical harm and content that is intended to suppress voting, we will remove the content.⁵

- 2. Downgrade the reach of all content from pages and groups that repeatedly post disinformation, so that it is seen by fewer users? If so, what is the expected timeline for this to take place? If not, why not?**

If a fact-checker rates a story as false, altered, or partly false, it will appear lower in News Feed. This significantly reduces the number of people who see it.⁶ Pages, Groups, websites, and Instagram accounts that repeatedly share content rated by fact-checkers as False or Altered will see their distribution reduced, and Pages and websites will see their ability to advertise removed.⁷

Groups, Pages, and accounts on Facebook and Instagram that repeatedly share false claims that violate our Community Standards, including claims about COVID-19 and vaccines that leading health organizations have debunked, face penalties and ultimately may be removed.

But it’s not enough to just limit misinformation that people might see. We also connect people to reliable information from trusted experts. We do this through centralized hubs like our COVID-19 Information Center, Climate Science Information Center, or US 2020 Voting Information Center, labels that we attach to certain posts with reliable information from experts, and notifications that we run in people’s feeds on both Facebook and Instagram.⁸

- 3. Suspend the advertising privileges of every advertiser that has repeatedly placed ads containing false and misleading information, whether or not the ads are currently**

³ <https://www.facebook.com/help/1952307158131536/>

⁴ <https://www.facebook.com/help/1952307158131536/>

⁵ <https://about.fb.com/news/2021/03/how-were-tackling-misinformation-across-our-apps/>

⁶ <https://www.facebook.com/help/1952307158131536/>

⁷ <https://www.facebook.com/help/1952307158131536/>

⁸ <https://about.fb.com/news/2021/03/how-were-tackling-misinformation-across-our-apps/>

active and whether or not the claim was debunked when the ad was placed? If so, what is the expected timeline for this to take place? If not, why not?

As explained in the response to your Question 2, pages and websites that repeatedly share content rated by fact-checkers as False or Altered will see their ability to advertise removed.⁹

Moreover, Facebook prohibits ads that include claims debunked by third-party fact-checkers or, in certain circumstances, claims debunked by organizations with particular expertise. Ads also must not violate our Community Standards (or ads on Instagram must not violate the Instagram Community Guidelines). Before ads show up on Facebook or Instagram, they're reviewed to make sure they meet our Advertising Policies. Typically most ads are reviewed within 24 hours, although in some cases it may take longer.¹⁰

You can find more information on our Advertising Policies relating to COVID-19 here: <https://www.facebook.com/business/help/1123969894625935>.

- 4. Disclose whether any of the disinformation referred to in the letter sent by me and several colleagues was amplified or accelerated by your curation algorithms and, if so, how and why that happened?**

Please see the response to your Question 2.

- 5. Commit to transparency regarding the standards governing decisions about whether or not to downgrade the algorithmic amplification of content determined to be disinformation by your fact-checking partners, or to downgrade and demonetize groups and pages that systematically spread disinformation? If so, what is the expected timeline for this to take place? If not, why not?**

Information about our approach to content rated by our fact-checking partners can be found in our Help Center.¹¹ Information about our policies regarding disinformation in ads can be found in our Advertising Policies.¹²

- 6. Publish reports detailing the actions taken pursuant to your policies on disinformation and misinformation, including the content, accounts, pages, and groups against which such actions were taken, and the criteria applied in deciding to take such actions? If so, what is the expected timeline for this to take place? If not, why not?**

We publish several reports related to content on our platforms that address many of these concerns.

Over the past three and a half years, we've shared our findings about coordinated inauthentic behavior we detect and remove from our platforms. As part of our regular CIB

⁹ <https://www.facebook.com/help/1952307158131536/>

¹⁰ <https://www.facebook.com/policies/ads/>

¹¹ <https://www.facebook.com/help/1952307158131536/>

¹² <https://www.facebook.com/policies/ads/>

reports, we're sharing information about all networks we take down over the course of a month to make it easier for people to see progress we're making in one place.¹³

We also publish our Community Standards Enforcement Report on a quarterly basis to more effectively track our progress and demonstrate our continued commitment to making Facebook and Instagram safe and inclusive. The following are policy areas we report on for Facebook: Adult Nudity and Sexual Activity, Bullying and Harassment, Child Nudity and Sexual Exploitation of Children, Dangerous Organizations: Terrorism, Dangerous Organizations: Organized Hate, Fake Accounts, Hate Speech, Regulated Goods: Drugs, Regulated Goods: Firearms, Spam, Suicide and Self-Injury, and Violent and Graphic Content.¹⁴

We also shared monthly reports for tackling COVID-19 misinformation with the House Energy & Commerce Committee throughout 2020.

7. Mr. Zuckerberg, social media platforms have been used to suppress the ability of many users to exercise their civil rights both online and offline. What systems have been implemented on your platforms to prevent this from happening? Do you have any data as to how successful these systems are?

Protecting and advancing the civil rights of our users is incredibly important to us. To that end, we voluntarily commissioned an independent civil rights audit in May 2018 to better understand the impact of our services on people's civil rights. The audit was a two-year-long investigation into our policies and practices, and it resulted in the creation of Facebook's Civil Rights team and the Civil Rights Task Force, both of which we've set up as long-term accountability structures that will continue to oversee the company's civil rights work.

The Civil Rights Team is led by Roy L. Austin, Jr., who brings with him over 25 years of experience working as both a civil rights lawyer and advocate. Roy is working to build out a team that will work to identify proactive civil rights priorities for the company and guide the implementation of those priorities; develop systems, processes, and other measures to improve the company's ability to spot and address potential civil rights implications in products and policies before they launch; and give voice to civil rights risks and concerns in interactions with leadership and executives, including the Civil Rights Task Force. The Team will also regularly meet with outside organizations, experts, and policymakers to understand potential civil rights issues.

The Civil Rights Task Force is something we set up in early 2019, in the midst of the civil rights audit. Its membership includes senior leaders at Facebook who are experts in areas such as content policy, fairness in artificial intelligence, and elections, along with key executives at the Company. The group currently meets quarterly for the purpose of surfacing, discussing, and addressing civil rights-related issues.

We have already implemented many of the recommendations set forth by our civil rights auditors, including:

¹³ <https://about.fb.com/news/2021/04/march-2021-coordinated-inauthentic-behavior-report/>

¹⁴ <https://transparency.facebook.com/community-standards-enforcement>

- Hate speech enforcement improvements;
- Building new tools to detect harassing behavior on the platform proactively, including detection of language that is harassing, hateful, or sexual in nature;
- Banning content that explicitly calls for harassment on our platform;
- Improving the transparency and consistency of our appeals decision-making;
- Updating our Community Standards to use a more expansive definition of national origin when applying our hate speech policies;
- Banning more than 250 white supremacist organizations from the platform by using a combination of AI and human expertise, including to remove content praising or supporting these organizations; and
- Elevating the role of Chief Diversity Officer to report directly to the Chief Operating Officer.

However, we know there is more to do and are committed to continuing to improve.

The Honorable Anna G. Eshoo

- 1. You testified that your users don't want to see misinformation or divisive content. Yet research shows that emotion-inducing content, including misinformation, is more engaging on social media. Facebook's business is based on engagement. How do you reconcile your testimony and what has been found by independent studies, such as the German Marshall Fund (GMF) report titled "New Study by Digital New Deal Finds Engagement with Deceptive Outlets Higher on Facebook Today Than Run-up to 2016 Election" (10/12/2020)?**

It is not in Facebook's interest—financially or reputationally—to push users towards extreme content. The company's long-term growth will be best served if people continue to use and value its products for years to come. If we prioritized trying to keep users online for a few extra minutes, but in doing so made them unhappy or angry and less likely to return in the future, it would be self-defeating. And the vast majority of Facebook's revenue comes from advertising. Advertisers don't want their brands and products displayed next to extreme or hateful content—a point that many made explicitly last summer during a high-profile boycott by a number of household-name brands. Even though troubling content is a very small proportion of the total (for example, hate speech is viewed 7 or 8 times for every 10,000 views of content on Facebook), the protest showed that Facebook's financial self-interest is to continue to reduce it, and certainly not to optimize for it.

People want to see accurate information on Facebook, and so do we. That's why we have made fighting misinformation and providing people with authoritative information a priority for the company. We have recalibrated our products and built global partnerships to combat misinformation on a massive scale.

We created an industry-leading fact-checking program. We work with more than 80 independent third-party fact-checking partners certified through the non-partisan International Fact-Checking Network to curb misinformation on Facebook and Instagram. If content is rated false by one of these third-party fact-checkers, we put a warning label on it. And based on one fact-check, we're able to kick off similarity detection methods that identify duplicates of debunked stories. When content is rated false, we significantly reduce its distribution; on average, this cuts future views by more than 80 percent. If people do try to share the content, we notify them of additional reporting, and we also notify people if content they have shared in the past is later rated false by a fact-checker. Group admins are notified each time a piece of content rated false by fact-checkers is posted in their Group, and they can see an overview of this in the Group Quality tool. We use information from fact-checkers to improve our technology so we can identify misinformation faster in the future.

We also work to reduce the incentives for people to share misinformation to begin with. Since a lot of the misinformation that spreads online is financially motivated spam, we focus on disrupting the business model behind it. We take action against Pages that repeatedly share or publish content rated false, including reducing their distribution and, if necessary, removing their ability to monetize. And we've enhanced our recidivism policies to make it harder to evade our enforcement. We've also taken steps to reduce clickbait and have updated our products so people see fewer posts and ads in News Feed that link to low-quality websites.

2. Have you empirically studied if Facebook and/or Instagram users show increased engagement with misinformation or divisive content? If so, what did you find about user engagement with misinformation or divisive content relative to other kinds of content?

We know that misinformation and divisive content can undermine the quality of public discourse and lead to polarization. In our case, it can also degrade the quality of our services. Our research suggests that no matter where we draw the line for what is allowed, as a piece of content gets close to that line, people will engage with it more on average—even when they tell us afterwards they don't like the content. That is why we've invested heavily in our integrity teams—now totaling about 35,000 people—and taken steps to try and minimize the amount of divisive news content people see in News Feed, including by reducing clickbait, engagement bait, and content rated as false by our independent fact-checkers.

We also frequently make changes to the algorithm that drives News Feed ranking in an effort to improve people's experience on Facebook. For example, in 2018, we responded to feedback from our community that public content—posts from businesses, brands, and media—was crowding out the personal moments that lead us to connect more with each other. As a result, we moved from focusing only on helping people find relevant content to helping them have more meaningful social interactions. This meant that people began seeing more content from their friends, family, and Groups. We also reduce the distribution of some problematic types of content, including content that users may find spammy or low-quality, such as clickbait headlines, misinformation as confirmed by third-party fact-checkers, and links to low-quality webpages like ad farms. When we did this, time spent on the platform actually dropped, but we believe that it increased meaningful engagement.

3. As I stated at the hearing, the Center for Countering Digital Hate found that the Explore and Suggested Posts parts of Instagram are littered with COVID-19 misinformation, election disinformation, and QAnon posts.

- a. How does Instagram select what appears in Explore and Suggested Posts?**
- b. Does Instagram recommend posts if a third-party fact-checker has debunked the post?**
- c. How does Instagram monitor images and videos for COVID-19 misinformation? What is the efficacy of any automated tools used to monitor for such misinformation?**

We want our users to trust what they see. Photo- and video-based misinformation is increasingly a challenge across our industry, and something our teams have been focused on addressing. We work with third-party fact-checkers in the US to help identify, review, and label false information. These partners independently assess false information on Facebook and Instagram to help us catch it and reduce its distribution. We have expanded our fact-checking program globally to allow fact-checking organizations around the world to assess and rate misinformation on our platforms.

On Instagram, we make recommendations to the people who use our services to help them discover new communities and content. Instagram may recommend content, accounts, and entities that people do not already follow through features like Explore, Accounts You May Like, and IGTV Discover. Our goal is to make recommendations that are relevant and valuable to each person who sees them. We do this by personalizing recommendations, which means making unique recommendations for each person. For example, if you interact with restaurants and bookstores on Instagram, we may recommend content about food, recipes, books, or reading.

When content on Instagram has been rated as False, Altered, or Partly False by a third-party fact-checker, we reduce its distribution by removing it from Explore and hashtag pages and reducing its visibility in Feed and Stories. In addition, it is labeled so people can better decide for themselves what to read, trust, and share, as is content a third-party fact-checker rates as Missing Context. When these labels are applied, they appear to everyone around the world viewing that content—in Feed, profiles, Stories, and direct messages.

We use image-matching technology to find further instances of this content and apply the label, helping reduce the spread of misinformation. In addition, if something is rated on Facebook, we automatically label identical content if it is posted on Instagram (and vice versa). The label will link out to the rating from the fact-checker and provide links to articles from credible sources that debunk the claim(s) made in the post. We also make content from accounts that repeatedly post content rated False or Altered harder to find by removing it from Explore and hashtag pages.

To determine which content should be sent to fact-checkers for review, we use a combination of feedback from our community and technology. We use “False Information” feedback reports and other signals to help identify and take action on potentially false information.

4. **A *New York Times* article titled “Facebook Struggles to Balance Civility and Growth” (11/24/2020) reported on you temporarily prioritizing “news ecosystem quality” (NEQ) to reduce misinformation.**
 - a. **Why did you *temporarily* increase the weight of NEQ in the recommendation algorithms?**
 - b. **Will you consider making this change permanent?**
 - c. **How did increasing the weight of NEQ impact user engagement across the platform?**
 - d. **How does Facebook calculate NEQ?**
 - e. **Has Instagram tested using NEQ or any equivalent metric for improving content quality?**

Facebook worked hard to do our part in helping to protect the integrity of the 2020 election, and we’re proud of the work we’ve done to support our democracy. Part of that work involved tackling misinformation and voter suppression. Sometimes the right thing to do from a

safety or security perspective isn't the best for privacy or free expression. The measure referenced in your question was a temporary change we made to help limit the spread of inaccurate claims about the election. News stories make up a small portion of what people see on Facebook overall, and political news makes up a smaller fraction of that news content. We're still ensuring that people see authoritative and informative news on Facebook, especially during major news cycles and around important global topics like elections, COVID-19, and climate change.

We believe that there is a lot to learn from the 2020 election, and we'll continue to study and learn valuable lessons from this election that will help us do even better in upcoming elections. We are in the process of evaluating all of our election-related measures to learn what we can from this experience, and to determine whether any of those measures should remain in place going forward.

5. Did American users of Facebook and Instagram see more or less election-related misinformation before the 2016 elections or before the 2020 elections? Please elaborate on how you determine this and what led to the change.

Facebook worked hard to do our part in protecting the integrity of the 2020 election, and we're proud of the work we've done to support our democracy. We know we have a responsibility to stop abuse and election misinformation on our platform, and we have made major investments since 2016 to better identify new threats, close vulnerabilities, and reduce the spread of viral misinformation and fake accounts.

We worked hard to tackle misinformation and voter suppression. Between March 1 and Election Day, we displayed warnings on more than 180 million pieces of content debunked by third-party fact-checkers that were viewed on Facebook by people in the US. We partnered with election officials to remove false claims about polling conditions, and we put in place strong voter suppression policies that prohibit explicit or implicit misrepresentations about how or when to vote, as well as attempts to use threats related to COVID-19 to scare people into not voting. We removed calls for people to engage in voter intimidation that used militarized language or suggested that the goal was to intimidate, exert control, or display power over election officials or voters. In addition, we blocked new political and issue ads during the final week of the campaign, as well as all political and issue ads after the polls closed on election night.

Since 2016, we've built an advanced system combining people and technology to review the billions of pieces of content that are posted to our platform every day. State-of-the-art AI systems flag content that may violate our policies, users report content to us they believe is questionable, and our own teams review content. We've also been building a parallel viral content review system to flag posts that may be going viral—no matter what type of content it is—as an additional safety net. This helps us catch content that our traditional systems may not pick up. We used this tool throughout this election, and in countries around the world, to detect and review Facebook and Instagram posts that were likely to go viral and took action if that content violated our policies.

6. You have touted your partnership with independent fact-checkers in your bid to avoid being an arbiter of truth. Yet you pay these groups, and so much so that some have directly told me they are dependent on Facebook for their existence. These

groups seem to be doing great work, but I'm worried about the systemic risk of this dependence. Are these organizations any more independent than credit rating agencies in the lead-up to the 2008 recession?

We rely on third-party fact-checkers certified through the non-partisan International Fact-Checking Network to help reduce the spread of false news and other types of viral misinformation, like memes or manipulated photos and videos, and we greatly value the independence these fact-checkers bring to their work. All of Facebook's third-party fact-checking partners are certified by Poynter's independent International Fact-Checking Network and must subscribe to the IFCN's rigorous Code of Principles.

IFCN signatories agree to commit to a standard of "non-partisanship and fairness" and agree to meet other journalistic standards like having a corrections policy and a policy of transparency of sources, methodology, and funding. IFCN launched its Code of Principles in 2016 to promote higher standards on accountability, transparency, and better signals in fact-checking. Fact-checking organizations applying to IFCN's Code go through a rigorous process supervised by independent assessors to get verified as a signatory by its board of advisors. Each signatory needs to go through a renewal process every year following the same steps.

7. If an account is banned from Facebook or Instagram for violating your Covid-19 or elections misinformation policies, does that ban apply to all Facebook-owned platforms? Why or why not?

We have strong policies against false claims about COVID-19 and vaccines that we enforce, and when a Page, Group, or person repeatedly violates these policies, we enforce penalties against them which can ultimately result in their removal. We don't automatically disable accounts across our apps, because the accounts may post about different things on our different services.

8. Why does your content moderation team report into your government affairs function?

a. Are moderation decisions independent of lobbying or campaign contribution decisions?

b. Are you aware of the leaders of your government affairs teams influencing content moderation decisions to aid the company's government affairs objectives?

Facebook is a platform for ideas across the political spectrum. We are committed to ensuring there is no bias in the work we do. Suppressing content on the basis of political viewpoints or preventing people from seeing what matters most to them is directly contrary to Facebook's mission and our business objectives.

Both our Content Policy team and the broader Public Policy team are guided by the same core principles, including the values of authenticity, safety, privacy, and dignity that are the foundation of our Community Standards.

9. Has Facebook ever taken adverse action on accounts of American users associated with religious minorities at the request of leaders of foreign countries?

We strive to enforce our policies consistently, without regard to religious or political affiliation. Our Community Standards apply to all content, and we assess everyone under those Standards. When we identify or learn of content that violates our policies, we remove that content regardless of who posted it.

Additionally, Facebook responds to government requests for data in accordance with applicable law and our terms of service. Each and every request we receive is carefully reviewed for legal sufficiency, and we may reject or require greater specificity on requests that appear overly broad or vague. As part of our ongoing effort to share more information about the requests we have received from governments around the world, Facebook regularly produces a report on government requests for user data to provide information on the nature and extent of these requests and the strict policies and processes we have in place to handle them. That report is available here: <https://govtrequests.facebook.com/government-data-requests>.

10. With regards to Spanish-language misinformation and Spanish-speaking content moderators, how does Facebook account for cultural differences of content moderators that come from diverse cultural experiences (e.g., Spain and South Texas)?

We work with over 80 independent third-party fact checkers certified through the non-partisan International Fact-Checking Network (IFCN). These fact-checkers review and rate content in more than 60 languages around the world. In the United States, we work with ten fact-checking partners, four of which fact-check content in both English and Spanish (Reuters, AFP, Factcheck.org and the Associated Press).

We have native Spanish-speaking content reviewers and fact-checking partners who are based in the US. Our machine learning technology is constantly developing as slang, colloquialisms, and other idioms are fed into it to improve accuracy.

We conduct Spanish-language content review 24 hours per day at multiple global sites. Spanish is one of the most common languages used on our platforms and is also one of the highest-resourced languages when it comes to content review.

Our reviewers do not work in isolation. There is quality control as well as management on site to help or seek guidance from if needed. We also have leadership, supervisors, trainers, and market specialists on site who can support our reviewers.

We have extensive training for our reviewers when they're onboarded and when there are clarifications or updates to policies. We have clear rules that are designed to eliminate bias and subjectivity and strict implementation standards and operational guidelines on enforcing our policies.

In developing our policies, we aim to be extremely prescriptive and try to write actionable policies that clearly distinguish between violating and non-violating content so as to make the decision-making process for reviewers as objective as possible. Our detailed

Community Standards are important and necessary, as they mitigate against subjectivity and bias in enforcement.

Every week, we audit a sample of every reviewer's decisions for accuracy and to ensure they are consistently enforcing our Community Standards. Where reviewers have made mistakes, we'll retrain individual reviewers. We also audit the auditors.

We have a comprehensive training program that includes at least 80 hours of live instructor-led training, as well as hands-on practice for all of our reviewers. Training can be thought of in three areas:

- **Pre-training:** In addition to receiving an introduction to Facebook, reviewers learn about what they can expect in the role. Part of this onboarding includes informing them of the resources available to them, such as resiliency and counseling as they begin work.
- **Hands-on Learning & Hands-on Classroom Instruction:** Over 80 hours of instructor-led training, practice and ongoing knowledge checks/quizzes. Following this in depth training we also ensure each reviewer spends time with veteran reviewers prior to working on live reports. This allows them time to get used to the tool and ask questions.
- **Ongoing Coaching & Training:** Routine coaching sessions, learning huddles, refresher sessions highlighted by areas for improvement, and ongoing policy update or clarification training.

All training materials are created in partnership with Content Policy and our in-market specialists or native speakers from the region. After starting, reviewers are regularly trained and tested with specific examples on how to uphold the Community Standards and take the correct action on a report. Additional training happens at a regular cadence, including when policies are clarified or as they evolve.

11. Given the violence in Myanmar, please provide the number of Burmese-speaking content moderators (employees and contractors) for each of Facebook, Instagram, and WhatsApp.

We have dedicated significant resources to the safety of our platform in Myanmar. Many teams at Facebook have worked over the past few years to better understand how our platform is used in Myanmar and how we can play a part in helping to prevent harm. We also built a team that is dedicated to Myanmar. This includes people who spend significant time on the ground working with civil society partners who are advocating on a range of human and digital rights issues across Myanmar's diverse, multi-ethnic society. Our goal is to understand and address current issues and those that are on the horizon.

Additionally, over the past several years, we have increased the number of Burmese and Sinhalese-language content reviewers as we continue to grow and invest in Myanmar and Sri Lanka. Our goal is always to have the right number of people with the right native language capabilities to ensure incoming reports are reviewed quickly and effectively. That said, there is

more to tackling this problem than reported content. A lot of abuse may go unreported, which is why we are supplementing our hiring with investments in technology and programs.

Following the military coup in Myanmar on February 1, the situation on the ground remains volatile, and Facebook is adapting to meet these events.

Our focus is threefold: First, do everything we can to prevent online content from being linked to offline harm and keep our community safe. Second, protect freedom of expression for the tens of millions of Myanmar citizens who rely on Facebook now more than ever. Third, ensure that Facebook, Messenger, and our family of apps stay online as a source of information and means of communication.

Facebook is treating the situation in Myanmar as an emergency. Our Integrity Operations Center has been running around the clock since the coup began. It brings together subject matter experts from across the company, including Myanmar nationals with native language skills, so we can monitor and respond to any threats in real time.

Beyond that, we've put several measures in place to support our community in Myanmar during this time.

On February 24, we announced that we would ban the remaining Myanmar military ("Tatmadaw") and military-controlled state and media entities from Facebook and Instagram, as well as ads from military-linked commercial entities. We have also indefinitely suspended the ability for Myanmar government agencies to send content removal requests to Facebook through our normal channels reserved for authorities around the world. Simultaneously, we are protecting content, including political speech, that allows the people of Myanmar to express themselves and to show the world what is transpiring inside their country. We're also taking the following additional steps:

- Introducing a new safety feature that allows users to lock their profiles so that non-friends cannot enlarge, share, or download their full-size profile picture or cover photo or see photos and posts on their timeline (both historic and new).
- Implementing a specific policy for Myanmar to remove praise, support, and advocacy of violence by Myanmar security forces and protestors from our platform.
- Continuing to enforce our policies on Coordinated Inauthentic Behavior (CIB) to combat influence operations when we find networks we previously removed trying to re-establish a presence on Facebook. We're doing this through a combination of automated and manual detection.
- Enforcing our Dangerous Individuals and Organizations policy to remove groups and individuals who seek to incite violence.
- Providing extra protections for journalists, civil society activists, human rights defenders, and deposed political leaders to prevent online threats against them,

and helping anyone who reasonably fears detention to secure their Facebook accounts and data from unauthorized access.

- Continuing to proactively remove content that violates our Community Standards, especially hate speech, incitement to violence, bullying and harassment, and misinformation that can lead to physical harm.
- Reducing the distribution of content in Myanmar that likely violates our hate speech and incitement policies—measures first taken during the November elections—as well as content that explicitly praises or supports the coup.
- Removing misinformation claiming that there was widespread fraud or foreign interference in Myanmar’s November election.
- Removing content that includes calls to bring weapons to any location across Myanmar.

We are committed to doing our part to keep people safe on our services and protect free expression, and we work hard to set and enforce policies that meet those goals. We will continue to invest extraordinary resources into content moderation, enforcement, and transparency.

12. On March 24th attorneys general from 12 states asked urged you to do more to combat COVID-19 misinformation. What changes will you make in response to the request?

Providing our users with authoritative information and addressing the spread of COVID-19 misinformation have been priorities for Facebook ever since the World Health Organization (“WHO”) declared COVID-19 a global public health emergency.

Facebook has dedicated significant resources to detecting misinformation—including false information about COVID-19 and vaccines—on our platform. Over the course of the pandemic, we have removed more than 16 million pieces of content from our apps for breaking our rules on COVID-19 and vaccine misinformation. We use machine learning to proactively identify and remove false information about COVID-19 and vaccines that could lead to imminent physical harm and is therefore in violation of our Community Standards. And we remove Groups, Pages, and accounts on Facebook that repeatedly share these debunked claims. We further use machine learning to identify other content that may be false and relay this content for review by our third-party fact-checkers. For claims about COVID-19 or vaccines that do not violate our Community Standards, but that have been identified as false by our third-party fact-checkers, we reduce the distribution of this content in News Feed, add a warning label, and disallow these claims in ads. Understanding that quick action is critical in keeping false claims from going viral, we take the extra step of temporarily demoting information that may be false in order to allow sufficient time for our fact-checkers to review it. We also use machine learning similarity detection methods to identify duplicates of debunked stories and reduce their distribution.

In recent months, we have continued to expand our efforts to remove false COVID-19 and vaccine-related information on Facebook. On February 8, in consultation with leading health

organizations, we expanded our list of false claims we remove to include widely debunked claims about COVID-19 and vaccines. Additionally, we take steps to reduce the distribution of vaccine content that is sensational or alarmist, criticizes someone's choice to receive or provide a vaccine, or promotes vaccine refusals, as well as shocking stories meant to discourage vaccination.

We work hard to keep our users informed with authoritative information about COVID-19 and vaccines. When COVID-19 misinformation is removed, users who may have liked, reacted, or commented on a post containing this content will receive a notification that says we've removed a post they've interacted with for violating our policy against COVID-19 misinformation that leads to imminent physical harm. Once a user clicks on that notification, they will see a thumbnail of the post and more information about where they saw it and how they engaged with it. They will also see information from authoritative sources explaining why it was false and why it was removed.

In addition, we have connected more than 2 billion Facebook and Instagram users to resources from health authorities like the WHO and Centers for Disease Control and Prevention ("CDC") through our COVID-19 Information Center and educational pop-ups, with over 600 million clicking through to learn more. These messages inform users about COVID-19 myths debunked by the WHO, including ones we've removed from our platform for leading to imminent physical harm. We also gave \$120 million in ad credits to help health ministries, NGOs, and UN agencies reach billions of people around the world with COVID-19 vaccine and preventive health information.

These are just some of the efforts that Facebook has taken to help keep users safe and informed during this global health crisis. Our efforts to address these challenging issues are ongoing, and we are continuing to work with health organizations to update and refine our approach to addressing COVID-19 and vaccine-related misinformation and promoting authoritative information about these topics.

- 13. A *New York Times* article titled "How Misinformation 'Superspreaders' Seed False Election Theories" (11/23/2020) found that a small number of users were responsible for spreading election disinformation on social media. How many of these accounts have you banned from your platform?**

Please see the response to your Question 5 above.

- 14. The Center for Countering Digital Hate recently published a report titled "The Disinformation Dozen" that identifies 12 individuals responsible for 65% of all anti-vaccine content on social media. For each of the 12 individuals, please identify the following, for each of Facebook and Instagram:**
- a. Have you taken adverse actions against any content the individual's account(s) posted? If so, how many times?**
 - b. Have you banned or taken other adverse actions against the individual's account(s)?**

We are aware of the twelve people whose accounts were identified by the Center for Countering Digital Hate as spreading false or misleading information about COVID-19 and vaccines. We have strong policies against false claims about COVID-19 and vaccines that we enforce. Pages, Groups, profiles, and Instagram accounts that repeatedly post misinformation or coordinate harm related to COVID-19, vaccines, and health may face restrictions, including (but not limited to) reduced distribution, removal from recommendations, or removal from our site.

We have reviewed the Facebook and Instagram accounts associated with the twelve people and have taken action against a number of them for violating our policies. On eleven separate occasions, individuals identified by the CCDH have had a presence on Facebook or Instagram permanently removed for violating our COVID-19 or vaccine misinformation policies. We will regularly update the claims we remove based on guidance from public health authorities as they learn more.

15. Have you observed a rise in hate directed at the Asian American and Pacific Islander (AAPI) on Facebook or Instagram in the prior year relative to the same 12-month period a year earlier?

a. If so, please share any quantifications of this change.

b. What have you done to combat anti-AAPI hate on your Facebook or Instagram?

Any hate or violence directed towards the Asian American and Pacific Islander community is abhorrent; we don't want any of this content on our platforms, and we work to remove it.

Our Hate Speech policy prohibits attacks against people on the basis of certain protected characteristics like race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity, and serious disease. Over the past year, we've worked hard to make sure we're applying this policy to the new types of speech we're seeing in connection with COVID-19. We remove claims that people of a certain protected characteristic have the virus, created it, or are trying to spread it. For example, the statements "Asians are bringing the virus into the US," "all Asians have COVID-19," or "Asians are responsible for creating the coronavirus" will be removed. And in advertisements, we remove terms like "Kung Flu," "China virus," and "Wuhan virus" when they're being used to sell merchandise or other goods. We assess speech trends, engage with experts, and provide additional policy guidance when appropriate. And to make sure the AAPI community feels safe and supported on our apps, we're expanding our comment warning features on Facebook to encourage people to reconsider harmful language and provide education around how that language may impact the AAPI community and others.

We are also prioritizing engagement with different communities, among them Asian American stakeholders, to share information about the steps we're taking to address hate speech and harassment and respond to their concerns. For example, we're hosting a series of Amplify API events throughout the week of May 17 to bring together business founders, community leaders, and media figures to discuss challenges the community is facing and how Asians and

Pacific Islanders are represented in media and culture, and to provide training and advice for API-owned small businesses. And given the recent rise in anti-Asian racism and the negative impact it has on mental health, we're partnering with organizations such as the Asian American Psychological Association, South Asian Americans Leading Together (SAALT), the Asian Mental Health Project and others to help people have conversations about mental health and encourage everyone to take a #MentalHealthMinute.

We're also expanding our investment in organizations committed to racial justice and equity. Last year, we invested \$10 million in racial justice grants, and we're expanding this investment with an additional \$5 million donation that will go to organizations focused on racial justice and equity for the Asian and Pacific Islander community and other marginalized communities.

16. What disinformation campaigns did you track in the lead up to the November 2020 elections that targeted people of color, women, LGBTQ+, veteran, or older American users?

- a. How did you respond to these campaigns?**
- b. Did you notify the users they had been targeted and share corrective information?**

For information about our efforts to fight election misinformation, please see the response to your Question 5 above. We also inform people when they encounter misinformation. We apply a label to content that's been reviewed by fact-checkers, so people can see additional context. And we notify people before they try to share this content, or if something they shared is later rated false. Group admins are also notified each time a piece of content rated false by fact-checkers is posted in their group, and they can see an overview of this in the Group Quality tool.

17. What categories of information related to the January 6th attacks have you shared with law enforcement agencies?

- a. Have you shared any of this information with researchers?**
- b. Have you permanently deleted any of this information?**
- c. Will you commit to retaining any such information until all legitimate law enforcement and research requests are fulfilled?**

We are committed to continuing to work with law enforcement and Congress as they investigate the attacks on the Capitol. As part of this cooperation, we are actively responding to valid legal process from federal and local law enforcement authorities, as well as proactively sharing some information when consistent with the law and our terms of service. And, in connection with these ongoing efforts, we are preserving the content related to valid law enforcement requests and the information we proactively share. We are not in a position to release details of our ongoing cooperation with law enforcement.

Even before the attacks on January 6, we had been focused on tracking potential extremist threats and engaging with law enforcement to provide information on them. For example, we were in regular contact with law enforcement authorities in Georgia leading up to the January 5 special election. On January 6 itself, our law enforcement response team was in regular contact with the FBI and Metropolitan Police Department, and we provided information related to multiple individuals who were livestreaming the attack on the Capitol.

Since January 6, our engagement with federal, state, and local law enforcement has of course continued. We have continued to send information proactively, and we have been responding to hundreds of legal requests—including subpoenas, warrants, and court orders. Throughout all of this, our law enforcement response team has spoken to dozens of federal, state, and local law enforcement agencies across the country, engaging daily with critical law enforcement points of contact.

18. What did you do to combat voter suppression targeted at Black Americans in the recent elections?

Attempts to interfere with or suppress voting undermine our core values as a company, and we work proactively to remove this type of harmful content. This was especially important in the midst of the COVID-19 pandemic to help deal with confusion about voting, voter registration, and election information ahead of the 2020 election.

Our Community Standards prohibit voter interference, and we worked with civil rights leaders and other experts to expand our policies significantly ahead of the 2020 election. For example, in addition to banning misrepresentations about the date, location, and time for voting or voter registration, we also banned attempts at coordinated interference that would affect an individual's ability to participate in an election and statements that voting participation would result in law enforcement consequences, both of which are often used to intimidate voters. A complete list of all voter interference policies is available here:

https://www.facebook.com/communitystandards/coordinating_harm_publicizing_crime.

For content that doesn't violate our Community Standards, we also adopted a policy of attaching a link to our Voting Information Center to all posts that discussed voting, including posts by politicians. Our Voting Information Center was designed to give millions of people accurate information about voting, while also giving them the tools they needed to register and make their voices heard at the ballot box. We worked with state election officials and other experts to ensure the Voting Information Center accurately reflected the latest information in each state. Through the use of this Voting Information Center, we ran the largest voting information campaign in American history. Based on conversion rates we calculated from a few states we partnered with, we estimate that we helped 4.5 million people register to vote across Facebook, Instagram, and Messenger and helped about 100,000 people sign up to be poll workers.

We had a dedicated team focused on preparing for the 2020 election. And since the most dangerous voter suppression campaigns can be local and run in the days immediately before an election, we also used our Elections Operations Center to quickly respond and remove false claims about polling conditions during the 72 hours before election day. Learning from our experience fighting COVID-19 misinformation, we partnered with and relied on state election

authorities to help determine the accuracy of information and whether it was potentially dangerous.

We will continue to work to protect the right to vote.

19. If a user’s post is designed to meet the letter of your policies but clearly infringes on the spirit of your misinformation policies, how do you treat the content and the content creator?

We work with over 80 independent third-party fact-checkers who are certified through the non-partisan International Fact-Checking Network (IFCN) to help identify and review misinformation. Content that is deemed by a fact-checker to be False, Altered, or Partly False, according to our public definitions, will have its distribution reduced, and it will appear lower in News Feed. We also implement an overlaid warning screen on top of fact-checked content. People who try to share the content will be notified of the fact-checker’s reporting and rating, and they will also be notified if content they have shared in the past has since been rated false by a fact-checker.

20. Spanish, Cantonese, Mandarin, Tagalog, Vietnamese, French and French Creole, Korean, German, Arabic, and Russian are the most spoken languages in the U.S. What are you doing to combat misinformation in these languages?

We work with over 80 independent third-party fact-checkers around the world, covering more than 60 languages, all of whom are certified through the non-partisan International Fact-Checking Network (IFCN) to help identify and review false news. Four of these partners work in both English and Spanish: Reuters, AFP, Factcheck.org, and the Associated Press. Additionally, if one of our fact-checking partners in France, for example, rates a post in French as False and a post with identical content is spreading in the US, we’ll label it as well.

If content is debunked by fact-checkers, its distribution will be reduced, and it will appear lower in News Feed. We also implement an overlaid warning screen on top of content marked as false. People who try to share the content will be notified of the fact-checker’s reporting and rating, and they will also be notified if content they have shared in the past has since been rated false by a fact-checker.

21. How often are you personally involved in content moderation decisions?

Mark Zuckerberg is generally not involved in day-to-day content moderation decisions, but he participates in conversations regarding particularly difficult or significant decisions.

22. How many employees and contractors at your company are dedicated to content moderation?

In the last few years, we have tripled, to more than 35,000, the number of people working on safety and security. This includes our growing team of more than 15,000 content reviewers—a mix of full-time employees, contractors, and companies we partner with. Our reviewers come from many backgrounds, reflect the diversity of our community, and bring a wide array of professional experiences, from veterans to former public sector workers.

a. How does this number compare to one year and four years prior?

Our content review teams around the world work 24 hours a day and in dozens of languages to review content. In 2017, we employed 10,000 individuals on our security and content review teams, a number that doubled—to a total of 20,000—by the end of 2018. Of that 20,000, approximately 7,500 were content reviewers.

In addition to our content review teams, to enforce our Community Standards, we have introduced tools that allow us to proactively detect and remove violating content using advances in technology, including artificial intelligence, machine learning, and computer vision. We do this by analyzing specific examples of bad content that have been reported and removed to identify patterns of behavior. Those patterns can be used to teach our software to proactively identify similar content. These advances in technology mean that we can now remove bad content more quickly, identify and review more potentially harmful content, and increase the capacity of our review team. To ensure the accuracy of these technologies, we constantly test and analyze our systems, technology, and AI to ensure accuracy. All content goes through some degree of automated review, and we use human reviewers to check some content that has been flagged by that automated review or reported by people that use Facebook. We also use human reviewers to perform reviews of content that was not flagged or reported by people to check the accuracy and efficiency of our automated review systems. The percentage of content that is reviewed by a human varies widely depending on the type and context of the content, and we don't target a specific percentage across all content on Facebook.

Our Community Standards Enforcement Report shares metrics on how Facebook is performing in preventing and removing content that violates our Community Standards. The report specifies how much content we took action on during the specified period, as well as how much of it we found before users reported it to us.

b. To what degree are your content moderation team members?

Our content moderators are a mix of full-time employees, contractors, and companies we partner with.

We have a comprehensive training program that includes many hours of live instructor-led training, as well as hands-on practice for all of our reviewers. In order to do their job well—and be well when doing it—people need good training. This means making sure our content reviewers have a strong grasp on our policies, the rationale behind them, and how to apply them. Training for all reviewers, including full-time employees, contractors, and partner companies, encompasses three areas:

- Pre-training, which includes what to expect on the job. Each hire also learns how to access resiliency and wellness resources, and gets information on how to connect with a psychologist when they need additional support.
- Hands-on Learning, including a minimum of 80 hours with a live instructor followed by hands-on practice using an actual replica of the system so new hires can practice in a “real” environment. Following hands-on training, reviewers get a

report highlighting areas where they're applying policies consistently and accurately and where they need more practice.

- Ongoing Coaching: Once hired, all reviewers get regular coaching, refresher sessions, and policy updates.

23. How many employees and contractors does your company employ for selling advertising?

The majority of our advertisers use our self-service ad platform to launch and manage their advertising campaigns. We also have a global sales force that is focused on attracting and retaining advertisers and providing support to them throughout the stages of the marketing cycle from pre-purchase decision-making to real-time optimizations to post-campaign analytics. We work directly with these advertisers, as well as through advertising agencies and resellers. We operate in more than 80 cities around the globe, the majority of which have a sales presence. We also invest in and rely on self-service tools to provide direct customer support to our users and partners.

24. To what degree are your product teams and executives incented – through bonuses, salary changes, or non-monetary incentives – to increase user engagement? Please describe any such incentives.

Facebook's mission is to bring communities closer together and to that end, we optimize for meaningful social interaction rather than for attention and time. We build our systems so that people see the most meaningful content from their friends, family, and Groups that they are part of. In fact, back in 2018, we announced a big shift in this direction knowing that people would spend less time on Facebook. Notably, we changed the way we approached News Feed rankings to focus not only on serving people the most relevant content, but also on helping them have more meaningful social interactions—primarily by doing more to prioritize content from friends, family, and Groups they are part of. We recognized that this shift would lead to people spending less time on Facebook, because Pages—where media entities, sports teams, politicians, and celebrities, among others, tend to have a presence—generally post more engaging (though less personally meaningful) content than a user's personal friends or family. The prediction proved correct; the change led to a decrease of 50 million hours' worth of time spent on Facebook per day, and we saw a loss of billions of dollars in the company's market cap. But we view this change as a success because it improved the experience of our users, and we think building good experiences is good for the business in the long term.

25. Have you studied the effectiveness of targeted or behavioral advertising relative to contextual advertising?

- a. What is the difference between targeted or behavioral advertising and contextual advertising with respect to user engagement and time spent on your platform?**
- b. Have you partnered with independent researchers to study this? If so, what did they find?**

We believe targeted advertising creates value for people and advertisers who use Facebook. Being able to target ads to the people most likely to be interested in the products, services, or causes being advertised enables businesses and other organizations to run effective campaigns at reasonable prices. This efficiency has particularly benefited small businesses, which make up the vast majority of the more than ten million active advertisers on Facebook.

26. As you have in the past, will you commit to continue promoting authoritative information to Facebook and Instagram users about Open Enrollment?

It is important, especially during a health crisis, for people to have health coverage and timely access to healthcare. We know that many people in the US do not have insurance, so we want to connect people to reliable information about the Affordable Care Act (ACA) and how to get coverage. Every year, there's an open enrollment period to purchase health insurance through the ACA. To raise awareness on this, we have launched a pop-up unit on Facebook and Instagram in the US, reminding and directing people to enroll. This follows a similar pop-up we launched at the start of the COVID-19 pandemic alerting people to the fact they could get insurance with ACA if they experienced a qualifying event (e.g., loss of a job). We are proud of these efforts and will continue to promote authoritative healthcare information to users.

The Honorable Yvette D. Clarke

Mr. Zuckerberg, in 2018, many fair housing groups and federal agencies alleged that Facebook violated the Fair Housing Act because its targeting systems allowed advertisers to exclude certain audiences by race, gender, and zip code. Facebook asserted that Section 230 immunizes them from civil rights on targeted ads.

- 1. While Facebook later settled the lawsuit and refined its internal advertising policies is this legal argument still Facebook’s position as a matter of law? Please explain.**

Facebook prohibits advertisers from discriminating against people based on personal attributes such as race, ethnicity, color, national origin, religion, age, sex, sexual orientation, gender identity, family status, disability, and medical or genetic conditions. Our Advertising Policies prohibit advertisers from using targeting options to discriminate against, harass, provoke, or disparage users, or to engage in predatory advertising practices.

Discrimination and discriminatory advertising have no place on Facebook’s platforms, and we remove such content as soon as we become aware of it. Our policies have long prohibited discrimination, and we have made significant changes to prevent advertisers from misusing our tools to discriminate in their ad targeting. As part of settlement agreements with civil rights organizations like the National Fair Housing Alliance, and based on ongoing input from civil rights experts, we have taken the industry lead by changing the way advertisers may select the audience for housing, employment, and credit (“HEC”) ads. Specifically, we have eliminated the ability to target HEC ads based on age, gender, or zip code, and we have severely restricted the number of interest category targeting options available. We’ve expanded our enforcement of these restrictions across all the tools businesses use to buy ads. Facebook does not have race data on its user base and has never given advertisers the ability to target ads based on race.

We’ve also added Housing, Employment, and Credit sections to the Ad Library, making it easy to search for and view all currently active US ads about housing, employment, and credit opportunities, regardless of the advertiser’s intended audience. We’ve also committed to studying the potential for algorithmic bias, including in our ad algorithms, with input from the civil rights community, industry experts, and academics.

We’re committed to preventing discriminatory advertising and are engaging in serious consultation with key civil rights groups, experts, and policymakers to help us find the right path forward.

- 2. Mr. Zuckerberg, what is your company doing to address the amplification and discrimination in ad-targeting?**

Please see the response to your previous question.

The Honorable Tony Cárdenas

1. What is the content moderation process for Spanish-language content in the United States?

We've been investing meaningfully in our machine learning over the past several years. Hopefully you have seen our external Community Standards Enforcement Report, where one critical metric that we track is how much of the violating content was detected by machine learning classifiers, versus how much was reported to us by humans. You can see in many of the charts that more and more of our content is detected via machine learning, which is critical for us in helping to identify violating content on the platform faster.

Generally, our machine learning is built by gathering what we call training data in local languages. Our classifiers run in a specific language, not on translated content. We also build our classifiers by using local language data. When a reviewer is making a decision on a piece of content, and they note that a piece of content violates our Community Standards, they don't just mark if the content is violating or not; they say what the violation was and, in many cases, get quite granular on the type of violation. This decision data becomes the training data we use to train our classifiers. This means our classifiers are constantly being retrained and improved based on real data, which is also helpful as content and language use evolves in the real world.

We're taking significant steps to fight the spread of misinformation using a three-part strategy:

- (1) we remove content that violates our Community Standards, including removing false claims about COVID-19 and vaccines that have been debunked by leading health organizations;
- (2) we reduce the distribution of content marked as false by independent third-party fact-checkers; and
- (3) we add warning labels that include additional context so people can decide what to read, trust, and share.

Our machine learning models to find potentially violating COVID-19 and vaccine content are trained to surface content in 19 languages using a data set of known violating content originally written in each of the 19 languages. For example, content originally written in Spanish is used to train the model to surface new Spanish content. We don't use any translations to train our health integrity models or those we use to surface content to our fact-checking partners, only original language data sets. We label data in the original language using labelers who speak those languages and train dedicated language models that don't translate any data.

In addition, we use machine learning to help identify potential misinformation that doesn't necessarily violate our Community Standards but is still eligible for fact-checking. We route this content to independent third-party fact-checking partners, and we may temporarily reduce this content's distribution pending review by a fact-checker.

If content is debunked by fact-checkers, its distribution will be reduced, and it will appear lower in News Feed. We also implement an overlaid warning screen on top of content marked as false. People who try to share the content will be notified of the fact-checker's reporting and rating, and they will also be notified if content they have shared in the past has since been rated false by a fact-checker.

Our machine learning and prioritization and detection systems include signals such as:

- **Disbelief comments:** Our system identifies when comments on a post include phrases that indicate people do not believe the content (e.g., “No way this is true!”). We have a set of disbelief comment signals that are in Spanish as well as other languages.
- **Fact-checkability:** Our system determines whether there is an actual claim that can be fact-checked as true or false. This is often referred to as a statement of fact, such as “Tomatoes are a fruit.”
- **Similarity detection:** Our system detects whether a piece of content is similar to content that a fact-checker in our program has already debunked.
- **User feedback:** Our system factors in whether someone in the community has marked the post as “false information.”

2. How is U.S. Spanish-language content routed to content moderators in scenarios where the content is in Spanish, but the user's Facebook settings are set to English (U.S.)?

We take steps to fight misinformation across our products in Spanish and dozens of other languages. We use the same machine learning model approaches in Spanish as we do in English to remove misinformation that violates our Community Standards. With respect to our third-party fact-checkers, we use similar techniques and signals in Spanish and English to surface content to fact-checking partners for them to consider reviewing.

Our enforcement against Spanish language content is two-pronged: we have Spanish-language content review for content that may violate our Community Standards; and we have partnerships with US-based third-party fact-checkers who review and rate content in Spanish.

Spanish-language content is reviewed by native speakers. The exceptions to this rule are infrequent, such as when a post has a significant amount of content from multiple languages.

3. Is the content presented to moderators for evaluation in its original Spanish, via automated translation to English, or both?

Content is provided to fact checkers in the language in which it is posted.

Spanish language content is reviewed by native speakers. The exceptions to this rule are infrequent, such as when a post has a significant amount of content from multiple languages.

- 4. Is the U.S. Spanish content routed to a specific team or content moderator with Spanish language fluency and specific cultural training?**

Please see the response to your Question 1.

- 5. How are algorithms trained and audited to ensure proactive detection, reporting, and processing of Spanish-language content in the U.S.?**

Please see the response to your Question 1.

- 6. How many U.S.-based Spanish-language content moderators does Facebook employ, as well as the overall number of U.S.-based content moderators? In each case, provide the breakdown between direct employees and contractors that are involved in moderation.**

Please see the response to your Question 1.

In the last few years, we have tripled the number of people working on safety and security and now have over 35,000, of whom about 15,000 are content reviewers who review content at more than 20 global sites in more than 50 languages. Annually, we spend more on safety and security than our whole revenue at the time of our IPO (more than \$3.7 billion), and the artificial intelligence required to help manage content at scale didn't exist until recently.

The Honorable Robin Kelly

- 1. At the hearing you said that you do not profit off harmful misinformation, conspiracy theories, and violent content.**
 - a. How do you manage to avoid collecting revenue from ads served on misinformation content?**
 - b. Are you claiming that your company has not received any payments from ads that displayed before or adjacent to harmful content and was later removed for violating your policies?**

Billions of people use Facebook and Instagram because they have good experiences; they don't want to see hateful content, our advertisers don't want to see it, and we don't want to see it. There is no incentive for us to do anything but remove it.

The Honorable Kathleen M. Rice

1. **Last year, Facebook rolled out the “Oversight board.” Their remit is limited to hearing appeals of Facebook content removals from users who are active on the platform, but the appeal must be filed within 15 days of the Facebook’s content decision. Why isn’t the Oversight Board given broader reach about the groups or pages that are allowed to remain on the platform?**

The board can already hear cases about content removed from and content left up on the platform. The board was originally launched with the ability to hear user-generated appeals on content we removed, as well as content decisions referred by Facebook itself, including content left up. Since launch, Facebook and the Oversight Board announced that, in addition to content taken down, users on Facebook and Instagram can appeal other people’s content that has been left up to the Oversight Board. The announcement represented a major expansion of the board’s scope, and we will continue to work to bring additional content into scope in the future.

While review of left-up content is an important piece of the oversight puzzle, there were a number of challenging technical and operational considerations we had to consider in the initial instance. Foremost, we had to guarantee that everything would be done effectively and in a privacy-conscious manner.

As outlined by the Oversight Board’s bylaws, Facebook’s goal is to make other content types available to the board for referral and review as quickly as possible. Other types of cases—for example, appealing actions taken to remove groups or Pages—have extra layers of complexity and will require additional time to build out and integrate with the board’s tooling. There is a growing team of engineers working on this effort, and we will continue to keep people updated on our progress to bring new types of content into scope for the board.

- a. **Why are the Oversight Board’s recommendations regarding Facebook’s underlying policies not binding? For example, a user cannot appeal the fact that a racist and extremist group is recommended to them or a decision by Facebook to not take action regarding racist or extremist content after being reported by that user. Why is this the case?**

We established the Oversight Board so people in the community can appeal our content decisions to a body that has independent judgment. The Oversight Board uses its independent judgment to decide some of our most significant and difficult cases, and the decisions it makes are binding.

The board’s policy recommendations are advisory, and we look to them for guidance in modifying and developing our policies and processes. If the Oversight Board issues a recommendation with its decision or makes a decision in response to a Facebook request for a policy guidance, Facebook is committed to reviewing the recommendations, and we will publicly respond within 30 days to explain how we will approach the recommendation.

The board’s recommendations are already having a demonstrable and significant impact on how we moderate content. We have responded to every single recommendation within 30

days, and we have committed to action on most of the board’s recommendations so far. We’ve made massive internal investments to respond to and implement the board’s recommendations—those changes guide how we moderate content for the almost 3 billion people on Facebook and Instagram.

But the board’s recommendations are just one input into our policy and operations processes. We have to balance what we hear from internal and external stakeholders, subject matter experts, and our community at large. As a result, we will not be able to implement every recommendation from the board.

Regarding the cases that can be appealed to the board, please see the response to your previous question.

2. Facebook’s newsfeed relies on algorithms that order content for users and recommend pages and groups to join. We know your algorithm is based on several variables such as “likes” and “shares.” When users “like” hateful content, is that activity then used to inform the recommendation algorithms that push content to those users?

On Facebook, people see posts from their friends, Pages they’ve chosen to follow, and Groups they’ve joined, among others, in their News Feed. On a given day, the number of eligible posts in a user’s Feed inventory can number in the thousands, so we use an algorithm to personalize how this content is organized. The goal of the News Feed algorithm is to predict what pieces of content are most relevant to the individual user, and rank (i.e., order) those pieces of content accordingly every time a user opens Facebook, to try and bring those posts that are the most relevant to a person closer to the top of their News Feed. This ranking process has four main elements: the available inventory (all of the available content from the people, Pages, and Groups a person has chosen to connect with); the signals, or data points, that can inform ranking decisions (e.g., who posted a particular piece of content); the predictions we make, including how likely we think a person is to comment on a story, share with a friend, etc.; and a relevancy score for each story, which informs its position in News Feed.

As part of this process, we reduce the distribution of many types of content—meaning that content appears lower in a person’s News Feed—for a variety of reasons, including because the content is sensational, misleading, gratuitously solicits engagement, or is found to be false by our independent fact-checking partners. For example, we demote clickbait (headlines that are misleading or exaggerated), highly sensational health claims (like those promoting “miracle cures”), and engagement bait (posts that explicitly ask people to engage with them).

Our approach goes beyond addressing sensational and misleading content on a post-by-post basis. When Pages or Groups repeatedly share some of these kinds of information on Facebook, like clickbait or misinformation, we’ll reduce the distribution of all of the posts from those Pages and Groups. And where websites generate an extremely disproportionate amount of their traffic from Facebook relative to the rest of the internet, which can be indicative of a pattern of posting more sensational and spammy content, we likewise demote all of the posts from the Pages run by those websites.

We've also adjusted other aspects of our approach to ranking, including fundamental aspects, in ways that would be likely to devalue sensational content.

Additionally, in 2018, we responded to feedback from our community that public content—posts from businesses, brands, and media—was crowding out the personal moments that lead us to connect more with each other. As a result, we moved from focusing only on helping people find relevant content to helping them have more meaningful social interactions. This meant that people began seeing more content from their friends, family, and Groups.

In addition to showing people content from the friends, Groups, and Pages they already follow in their News Feed, we make recommendations to help them discover new things we think they are likely to be interested in, across surfaces like Groups You Should Join and Pages You Might Like. But since recommended content doesn't come from accounts people choose to follow, we have more stringent standards for what we recommend, and we follow our Recommendation Guidelines to determine what content is eligible for recommendation. For example, our goal is to not recommend content that is false or misleading or that is shared by Groups or Pages that have recently violated our policies.

- 3. One year ago, in March 2020 Facebook announced that, in light of the pandemic, outside of a small team of content reviewers allowed to work from home, the vast majority of your 35,000 person content review team was no longer reviewing content. Can you tell us roughly how many content reviewers are currently operational? And how has this impacted Facebook's ability to moderate violative content?**

In the last few years, we have tripled—to more than 35,000—the number of people working on safety and security. This includes our growing team of more than 15,000 content reviewers—a mix of full-time employees, contractors, and companies we partner with. Our reviewers come from many backgrounds, reflect the diversity of our community, and bring a wide array of professional experiences, from veterans to former public sector workers.

- 4. According to the ADL's latest survey, only 14% of individuals who reported a physical threat to a social media platform said the platform deleted the threatening content—a significant drop from 22% the prior year. We have heard you say that the scale of your platform impacts your ability to moderate content effectively. Why aren't you investing more into content moderation and protecting targets of harassment?**

Our Community Standards Enforcement Report details the action we've taken on violating content. It is available here: <https://transparency.facebook.com/community-standards-enforcement>. With respect to bullying and harassment, content actioned increased from 3.5 million pieces of content in the third quarter of 2020 to 6.3 million in the fourth quarter of 2020. This was driven by increasing our automation abilities and improving our technology to detect and remove more English-language comments, which helped our proactive rate increase from 26.4% to 48.8%, in addition to regaining some manual review capacity in the third quarter.

We have invested significantly in the area of safety and security. We have over 35,000 people working on safety and security across our platforms, about 15,000 of whom review

content. We're spending as much—if not more—on safety and security as the entire revenue of our company at the time of our IPO.

We recognize how important it is for Facebook to be a place where people feel empowered to communicate, and we take our role in keeping abuse off our service seriously. We do not tolerate harassment on Facebook because we want people to feel safe to engage and connect with their community. Our harassment policy applies to both public and private individuals and includes behavior like repeatedly contacting a single user despite that person's clear desire and action to prevent that contact and repeatedly contacting large numbers of people with no prior solicitation. It also applies to calls for death, serious disease, disability, or physical harm aimed at an individual or group of individuals in a message thread. Context and intent matter, however, and we allow people to share and re-share posts if it is clear that something was shared in order to condemn or draw attention to harassment. The detailed guidelines our reviewers use to assess whether content violates our hate speech policies are available at <https://www.facebook.com/communitystandards/safety/harassment>.

5. After nearly a decade of advocacy by the ADL and other Jewish-led organizations, Facebook announced that it changed how its platforms categorized Holocaust denial content—that it would no longer be classified as misinformation, and instead hate speech. What are the different content moderation methods you use for something that is classified as misinformation versus disinformation (intentional spreading of false information) versus hate speech?

a. While we applaud you for finally changing the policy around Holocaust Denial content, a recent study done by ADL found significant amounts of Holocaust Denial still pervading the platform. How are we to hold you accountable for the egregious enforcement lapses occurring on your platform?

Anti-Semitism is abhorrent and has no place on our platform. For years, Facebook has removed any post that celebrates, defends, or attempts to justify the Holocaust. The same goes for any content that mocks Holocaust victims, accuses victims of lying about the atrocities, or advocates for violence against Jewish people in any way.

Last year, we took another step in our effort to fight anti-Semitism and hate on our services by updating our hate speech policy to prohibit any content that denies or distorts the Holocaust.

We take a zero-tolerance approach to hate speech we find on our platform. We invest billions of dollars each year in people and technology to keep our platform safe and are leading the way when it comes to deploying artificial intelligence technology to remove hate speech at scale. We have over 35,000 people working on safety and security, about 15,000 of whom review content. The European Commission's recent assessment on the Code of Conduct on Countering Illegal Hate Speech Online found that Facebook assessed 95.7 percent of hate speech reports in less than 24 hours; this number jumped to 99.1 percent over a 48-hour period.

In February, we published our Community Standards Enforcement Report for the fourth quarter of 2020, which includes information on our hate speech enforcement and demonstrates

our commitment to these efforts. In the fourth quarter of 2020, we took action on 26.9 million pieces of content, more than 97% of which we acted on before any user reported it. We also disclosed how prevalent hate speech violations were on Facebook in that quarter, noting that of every 10,000 content views, our estimates indicate that only 7 or 8 of those views contained hate speech. We've come a long way, and we know we still have a ways to go. With the use of prevalence, user reports, AI, and a combination of our internal teams and third-party experts, we're working to keep hate speech off our apps.

With respect to misinformation and disinformation, we generally use "misinformation" to mean the sharing of false content without intent to manipulate, and we generally use the term "disinformation" for the sharing of information with the intent to manipulate. Our work relating to misinformation and disinformation is reflected in our content moderation efforts, misinformation policies, third-party fact-checking program, and efforts to fight coordinated inauthentic behavior.

When it comes to misinformation, we work with over 80 independent third-party fact-checkers who are certified through the non-partisan International Fact-Checking Network (IFCN) to help identify and review false news. If content is deemed by a fact-checker to be false, altered media, or partly false, its distribution will be reduced, and it will appear lower in News Feed. We also implement a warning screen on this content, as well as content a fact-checker rates as missing context. People who try to share the content will be notified of the fact-checker's reporting and rating, and they will also be notified if content they have shared in the past has since been rated by a fact-checker.

When it comes to disinformation, we are looking at specific coordinated inauthentic behaviors exhibited by violating actors, and we are actively working to find and stop coordinated campaigns that seek to manipulate public debate across our platforms. We have grown our team focused on disrupting influence operations to over 200 experts across the company, with backgrounds in law enforcement, national security, investigative journalism, cybersecurity, law, and engineering. We continue to expand our technical teams to build scaled solutions to help detect and prevent these behaviors, and we are partnering with civil society organizations, researchers, and governments to strengthen our defenses. We have also improved our detection systems to more effectively identify and block fake accounts, which are the source of a lot of the inauthentic activity.

Over the past three and a half years, we've also publicly shared our findings about coordinated inauthentic behavior we detect and remove from our platforms. As part of our regular CIB reports, we're sharing information about all networks we take down over the course of a month to make it easier for people to see progress we're making in one place. All of our CIB reports may be found here: <https://about.fb.com/news/tag/coordinated-inauthentic-behavior/>.

- 6. There have been some successful partnerships between big tech companies like yours and law enforcement. However, the sales of illicit live animals and their parts and products continue to grow in online marketplaces and communities of violence promoting illegal animal abuse content continue to be prevalent problems. Some online crime experts have raised concerns that the widespread company policies of removing illicit wildlife for sale may be hindering law enforcement efforts by prematurely alerting criminals who then open new accounts under different**

usernames. It also poses the problem of deleting public evidence that could help authorities catch traffickers. Company policies also differ between platforms, creating an inconsistent landscape for both companies and law enforcement to work in.

a. How could Congressional mandates assist your company in building stronger and more comprehensive partnerships with law enforcement?

Animal abuse is abhorrent, and we have no tolerance for any content promoting it. We also work hard to prevent the sale of illicit live animals, as well as illicit animal parts or products, on Facebook Marketplace.

Facebook and Instagram have been working with the World Wildlife Fund (WWF) since 2016 to better understand the global landscape of wildlife trafficking and the role that social media platforms inadvertently play in facilitation. We regularly participate in annual Interpol Wildlife Crime Working Group meetings that bring together law enforcement, government officials, and NGOs to address wildlife crime issues. We also train law enforcement from different countries on ways to address, report, and send us lawful process on such matters.

We welcome the opportunity to work together with Congress on how we can better address these issues and build even stronger partnerships with law enforcement in this space.

The Honorable Lori Trahan

- 1. Children who use YouTube Kids often migrate to the main YouTube platform.¹⁵ According to Facebook’s internal metrics, has the increased usage of Kids Messenger led to an increase in monthly active users—predicted to be under the age of 13—within the Facebook ecosystem (WhatsApp, Facebook or Instagram)?**

When a parent signs their child up for a Messenger Kids account, we don’t create a profile for the child on Facebook, and we don’t automatically migrate their account to Facebook when they turn 13.

We require everyone to be at least 13 years of age to use Facebook and Instagram. When we learn an underage user has created an account, we remove that user from our platforms.

We also take a continuous, multi-layered approach to understanding and refining the accuracy of the age of our users. Effectively verifying age should not focus on a single-step process, but rather a collection of ongoing efforts that work dynamically to provide effective solutions.

Our multi-layered approach starts with requiring users to provide their date of birth when they register new accounts. This is done through an age-neutral process with technical restrictions to make it harder for users to provide false information. We also allow anyone—whether they use our services or not—to report suspected underage users on Instagram and Facebook. We have dedicated channels where we review these accounts based on these or other reports. While this approach has been appropriate in light of industry standards to date, we continue to invest in new solutions that provide ongoing age assurances.

We continue to develop and refine AI tools to identify users under 18. These tools add another layer of confidence to user-stated ages, allowing us to provide age-appropriate experiences. It’s also important that we make sure these tools are sufficiently accurate before using them to limit or restrict people’s use of our services. We’re committed to making progress with this technology, but it will take time to get this right.

One of the greatest challenges our industry faces is ensuring that users below the minimum age do not gain access to services that are not designed for them. Parents should not have to choose between allowing their children on apps that weren’t designed for them or restricting their access to apps entirely. While the issue of children under 13 misrepresenting their age will continue to be an industry-wide problem, apps designed specifically for kids can give parents more control over options and reduce the incentive for kids to be dishonest about their age.

Closer industry collaboration is also critical to developing effective and scalable measures to ensure young people consistently receive age-appropriate experiences across the online ecosystem. We believe strongly that providers across our industry must come together

¹⁵ Ryan Mac and Craig Silverman, *Facebook Is Building An Instagram For Kids Under The Age of 13* (March, 2021), BuzzFeed, www.buzzfeednews.com/article/ryanmac/facebook-instagram-for-children-under-13.

with governments and experts to arrive at thoughtful solutions designed with the best interests of young people in mind.

2. **In 2017, Facebook’s leaked internal documents which described that the company “can monitor posts and photos in real time to determine when young people feel “stressed”, “defeated”, “overwhelmed”, “anxious”, “nervous”, “stupid”, “silly”, “useless” and a “failure.”¹⁶ Facebook later announced new guidelines for research at the company, including “[i]f proposed work is focused on studying particular groups or populations (such as people of a certain age) or if it relates to content that may be considered deeply personal (such as emotions) it will go through an enhanced review process before research can begin.”¹⁷**
 - a. **Since 2017, how many projects have undergone “enhanced review”?**
 - b. **Has the “enhanced review” process been audited by an independent third party for compliance?**

As we explained at the time, we do not offer tools to make recommendations to people based on their emotional state. The 2017 research referenced above did not follow the established processes we have to review the research we perform. The research was also never used to target ads and was based on data that was anonymous and aggregated.

Facebook does research in a variety of fields, from systems infrastructure to user experience to artificial intelligence to social science. We do this work to understand what we should build and how we should build it, with the goal of improving the products and services we make available each day. We’re committed to doing research to make Facebook better, but we want to do it in the most responsible way.

In October 2014, we announced a framework that covers both internal work and research that might be published. We have continued to improve upon this framework, including talking about changes (<https://research.fb.com/blog/2016/06/research-review-at-facebook/>) and publishing a law review article (<https://scholarlycommons.law.wlu.edu/wlulr-online/vol72/iss3/8/>) detailing how our research process worked. The main elements are:

- **Training:** All new employees learn about Facebook’s policies around data access and privacy. Anyone who conducts research also attends a multi-week bootcamp where they learn specifically about our research review program, why it matters, and how it works. We’ve incorporated education on our research practices into Facebook’s six-week training program, called bootcamp, that new engineers go through, as well as training for others doing research. We’ll also include a section on research in the annual privacy and security training that is required of everyone at Facebook.

¹⁶ Sam Levin, *Facebook told advertisers it can identify teens feeling ‘insecure’ and ‘worthless,’* (May, 2017), The Guardian, www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens.

¹⁷ Facebook, *Research at Facebook*, (October, 2014), Facebook, about.fb.com/news/2014/10/research-at-facebook/.

- **Guidelines:** We've given researchers clear guidelines. If proposed work is focused on studying particular groups or populations or if it relates to content that may be considered deeply personal, it will go through an enhanced review process before research can begin. The guidelines also require further review if the work involves a collaboration with someone in the academic community.
- **Review:** All research must be approved by a research lead trained in data ethics. Research leads refer more sensitive projects to the enhanced review process. Similarly, any research that falls within our guidelines similarly automatically goes through the enhanced review process. We've created a panel including our most senior subject-area researchers, along with people from our engineering, research, legal, privacy, and policy teams, that will review projects falling within these guidelines. This is in addition to our existing privacy cross-functional review for products and research.
- **Research Website:** Our published academic research is now available at a single location (<https://research.facebook.com/>) and is updated regularly.

3. Researchers have found that Facebook's ad targeting algorithm is discriminatory and reinforces filter bubbles.¹⁸ The full impact of this "skewing" is difficult for the public to monitor in part because Facebook's Ad Library is missing key pieces of data.¹⁹

- a. Will Facebook consider adding data related to an ad's engagement (e.g., number of likes, shares, and video views an ad received) to the Facebook Ad Library? If not, why not?**

Engagement data (i.e., likes, comments) represents a very small percentage of the people who actually saw an ad. We believe that metrics like reach (i.e., impressions) and spend paint a much better picture of the impact that an ad had, which is why we include those in the Ad Library for political ads.

- b. Will Facebook consider adding data related to how an ad was targeted to the Facebook Ad Library? If not, why not?**

We have heard the request from the researcher community for better understanding of targeting for social issue, electoral, and political ads featured on Facebook and Instagram during the past election. That's why we made this data package available to researchers who apply to participate in the Facebook Open Research & Transparency (FORT) platform (<https://research.fb.com/data/>). The data package contains social issue, political, and electoral ads

¹⁸ Ali, M., Sapiezynski, P., Korolova, A., Mislove, A. and Rieke, A., 2019. Ad delivery algorithms: the hidden arbiters of political messaging. *arXiv preprint arXiv:1912.04255*. see also Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A. and Rieke, A., 2019. Discrimination through optimization: How Facebook's Ad delivery can lead to biased outcomes. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), pp.1-30.

¹⁹ UpTurn, *Leveling the Platform: Real Transparency for Paid Messages on Facebook*, (May 2018), <https://www.upturn.org/reports/2018/facebook-ads/>.

(approximately 1.65 million ads) that ran in the US during the 90 days before the US 2020 elections, from August 3 to November 3, 2020. Targeting information is provided in a secure environment specially created to enable academic researchers to study the impact of Facebook's products, with a particular focus on elections and democracy, while protecting people's privacy.

We agree that greater ads transparency is important, but it can't be at the cost of people's privacy. FORT's secure environment is a first step toward learning how we can help academic researchers study the impact of Facebook's products while protecting people's privacy. We will continue to learn and to work to provide academics and researchers with the information they need to study the impact of digital ads on elections.

c. Will Facebook consider adding data related to the aggregated demographic attributes, including race and political affiliation, of an audience that an ad was delivered to the Facebook Ad Library? If not, why not?

Facebook is committed to transparency for all ads, including ads with political content. That's why we've endorsed legislation like the Honest Ads Act, which aims to improve ad disclosures.

Facebook believes that people should be able to easily understand why they are seeing ads, who paid for them, and what other ads those advertisers are running. Our Ad Library is a unique tool that allows the public to see every ad served to anyone in an easily searchable database. We are continually investing in and considering various ways to improve the Ad Library.

Last year, for example, we announced changes to provide more transparency over who is using ads to try to influence voters and to give people more control over the ads they see:

- View audience size in the Ad Library: We've added ranges for Potential Reach, which is the estimated target audience size for each political, electoral, or social issue ad, so you can see how many people an advertiser wanted to reach with every ad.
- Better Ad Library search and filtering: We've added the ability to search for ads with exact phrases, better grouping of similar ads, and adding several new filters to better analyze results—e.g., audience size, dates, and regions reached. This allows for more efficient and effective research for voters, academics, or journalists using these features.
- Control over Custom Audiences from a list: We rolled out a control to let people choose how an advertiser can reach them with a Custom Audience from a list. These Custom Audiences are built when an advertiser uploads a hashed list of people's information, such as emails or phone numbers, to help target ads. This control is available to all people on Facebook and applies to all advertisers, not just those running political or social issue ads. People have always been able to hide all ads from a specific advertiser in their Ad Preferences or directly in an ad. But now they are able to stop seeing ads based on an advertiser's Custom

Audience from a list—or make themselves eligible to see ads if an advertiser used a list to exclude them.

- See fewer political ads: Seeing fewer political and social issue ads is a common request we hear from people. That’s why we added a new control that will allow people to see fewer political and social issue ads on Facebook and Instagram. This feature builds on other controls in Ad Preferences we’ve released in the past, like allowing people to see fewer ads about certain topics or remove interests.

4. In 2020, Facebook announced that it was going to reclassify news pages with “direct and meaningful ties” to political organizations as political advertising, removing the advantages associated with being labeled as “news.”²⁰ However, researchers at the Tow Center have found that Facebook’s labeling is not always correct and is difficult to monitor.²¹ Will Facebook consider making an API available to researchers and the public to see what is indexed as a “Media/News Company?” If not, why not?

Facebook values transparency and has established a clear set of guidelines around news Pages and what entities are considered to have direct, meaningful ties to political entities.

As an initial matter, the news Page index is a foundational step to help us identify Facebook pages that primarily publish news content on the platform. Any Facebook Page that primarily creates journalism should be registered as a news Page on Facebook. In order to be eligible for inclusion in the news Page index, Pages must be active in the last 90 days, must have existed for at least 90 days, and must link to a functioning website with a verified domain. The publisher’s Page name must reflect a relationship to its associated website or owning business. Moreover, the information provided by the Publisher in the “About” section of their Page may be assessed for accuracy and whether it reflects their status as a News provider with a multi-person editorial staff. Additionally, the majority of links posted on the publisher’s Page must come from the website verified for news Page registration.

Based on industry standards, we expect publishers to submit news Pages that:

- Primarily create journalism that reports on current events or timely information, and that is not parody or satire.
- Cite and/or link to fact-based sources for published information. Publishers that cite parody and satire as fact, or that repeatedly cite known publishers of misinformation and hoaxes, will not meet this requirement.
- Include dates/timestamps on published content.

²⁰ Sara Fischer, *Exclusive: Facebook cracks down on political content disguised as local news*, (August, 2020), www.axios.com/facebook-pages-news-exemption-e66d92ce-2abd-4293-b2ad-16cf223e12f1.html.

²¹ Emily Bell and Sara Sheridan, *Google and Facebook Have a News Labeling Problem*, Columbia Journalism Review (October, 2020), www.cjr.org/analysis/google-and-facebook-have-a-news-labeling-problem.php.

- Primarily publish content that is not user-generated or aggregated from other websites.
- Provide transparent information about writers or editors (e.g., full names of writers in bios, or of writers and editors in a staff directory).

Registered news Pages based in the United States that Facebook believes to have direct, meaningful ties to political entities are subject to a policy that limits certain features that are available to registered publishers. Examples of features that these Pages won't have access to include, but may not be limited to, exemption from the authorization and disclaimer process (<https://www.facebook.com/business/help/387111852028957>) when running ads about social issues, elections, or politics; the Page will not be eligible for inclusion in Facebook News (<https://www.facebook.com/news>); and will not have access to news messaging on the Messenger Business Platform, or the WhatsApp Business API. This impacts Pages that meet any of the following criteria:

- They are owned by a Political Entity or a Political Person;
- The news publisher or its owning entity has a Political Person leading (1) the company (e.g., the Chief Executive Officer), (2) the board of directors (e.g., Chairman of the Board), or (3) the editorial function of the publisher, (e.g., the editor-in-chief);
- The news publisher shares proprietary information about any of its Facebook, Inc. accounts or account passwords, API access keys, and/or data about their Facebook readers (e.g., location, demographics, consumption habits) directly with a Political Person or Entity; or
- The Page lists a Political Entity or a Political Person as its Confirmed Page Owner or Confirmed Page Partner pursuant to our Confirmed Page Owner Policy.

A "Political Entity" is an organization, company, or other group whose primary purpose is to influence politics, public policy, and/or elections. This includes:

- Political parties, campaigns for elected office, ballot initiative campaigns, entities regulated under Section 527 of the Internal Revenue Code ("IRC") (e.g., Political Action Committees and Super Political Action Committees), and some entities regulated as "Social Welfare Organizations" under Section 501(c)(4) of the IRC that can or do primarily engage in efforts to influence public policy or elections are Political Entities.
- For-profit businesses that provide political consulting or strategic communications services to other types of Political Entities listed above (e.g., campaigns), will also be considered Political Entities themselves.

A "Political Person" is a candidate for elected office, a person who holds elected office, a person whose job is subject to legislative confirmation, or a person employed by and/or vested with decision-making authority by a Political Person or at a Political Entity.

The Honorable Robert E. Latta

Mr. Zuckerberg, you created the Facebook Oversight Board in order to provide some accountability to Facebook's actions. That board is supposed to be independent. Please provide answers to the below questions regarding the oversight board:

1. Are the members of the Board paid? If so, by whom?

The Oversight Board Trust controls the finances for the board and the LLC that oversees the contracts with members and employment relationships with board staff. We released the associated documents for the trust and LLC in December 2019. Facebook has committed and transferred to the trust \$130 million to fund the board's operations. This funding is irrevocable, meaning the board is already ensured funding for two full three-year terms. Member compensation is set by the trustees and is issued on a schedule based on the fulfillment of duties. It is not conditioned or withheld based on the outcome of board decisions.

2. How are the members appointed?

To support the initial formation of the board's first twenty members, Facebook selected four co-chairs from a slate of qualified candidates. The co-chairs and Facebook then jointly worked together to select, interview, and make offers to candidates to fill the remaining board positions. All members were formally appointed by the trustees. Our goal was to assemble a board with a diverse set of perspectives; a proven track record of deliberating thoughtfully and collegially and protecting freedom of expression; and significant expertise in digital rights, religious freedom, content moderation, digital copyright, online safety, internet censorship, platform transparency, and civil rights.

Facebook will work with the current members and co-chairs of the Oversight Board to select the remaining 20 members. After that point, the Oversight Board's membership committee will take on responsibility for selecting future members, independent of Facebook.

Earlier this year, after one of the original 20 board members stepped down, the Oversight Board oversaw the naming of their replacement without input from Facebook—the replacement of departing members is something the board is solely responsible for.

3. How long do the members serve?

Each member serves for a three-year term, for a maximum of three terms. Trustees oversee the approval of term renewals.

4. Can they be removed? If so, by whom?

The Oversight Board is comprised of three interlocking elements: board members, the Trust, and the Administration. The roles and responsibilities of these three elements ensure mutual accountability in realizing the board's purpose. Trustees are responsible for safeguarding the independence of the board and for ensuring that the board operates effectively in fulfilling its stated purpose.

A member may only be removed prior to the expiration of their term if the trustees have determined that that member has violated the code of conduct. Members will not be removed due to content decisions they have made. The individual trustees may receive, verify, and act upon requests to remove members based on violations of the code of conduct. These requests may come from the board, the director, or the public. Removals require a two-thirds vote of the board (not counting the member(s) in question), subject to approval of the trustees.

5. What free speech principles do they follow? Does this include the first amendment to the United States Constitution?

A commitment to freedom of expression was a key criterion for Facebook in selecting members for the board, and every one of the board's members had written, ruled, advocated, or made decisions in ways that made their commitment to free speech clear when they were selected. The board members also have agreed to operate in accordance with the Oversight Board charter, which begins with the statement that "Freedom of expression is a fundamental human right."

The Oversight Board determines whether enforcement decisions were consistent with Facebook and Instagram content policies and values, as well as Facebook's commitment to upholding freedom of expression within the framework of international norms of human rights. These content policy values are listed here, with descriptions quoted from Facebook's Community Standards:

- Voice: "The goal of our Community Standards has always been to create a place for expression and give people voice. . . . Building community and bringing the world closer together depends on people's ability to share diverse views, experiences, ideas and information."
- Authenticity: "We want to make sure the content people are seeing on Facebook is authentic. We believe that authenticity creates a better environment for sharing, and that's why we don't want people using Facebook to misrepresent who they are or what they're doing."
- Safety: "We are committed to making Facebook a safe place. Expression that threatens people has the potential to intimidate, exclude or silence others and isn't allowed on Facebook."
- Privacy: "We are committed to protecting personal privacy and information. Privacy gives people the freedom to be themselves, and to choose how and when to share on Facebook and to connect more easily."
- Dignity: "We believe that all people are equal in dignity and rights. We expect that people will respect the dignity of others and not harass or degrade others."

6. How many cases does the Board take up?

The Oversight Board reviews a select number of highly emblematic cases and determines if Facebook's decisions were made in accordance with its stated values. There is no specific

number or cadence mandated by the bylaws; case selection has been an ongoing process at the sole discretion of the board since the board first announced it was accepting case submissions in October 2020. The board chooses the cases it wants to hear among the eligible cases referred to it either by user or reporter appeal, or from Facebook. Certain types of content may not be eligible for user appeals because they are not technically or operationally feasible to be sent to the board, or would result in a violation of law by Facebook. The only exception to the board solely choosing the cases it hears will be cases referred to the board by Facebook using expedited review. In exceptional circumstances, including when content could result in urgent real-world consequences, Facebook may send cases for an automatic and expedited review, which will be completed within a maximum of thirty days.

7. What is the time frame to hear and resolve an appeal?

Save in exceptional circumstances, the timeframe for case decisions and implementation will be a maximum of 90 days, starting from the date the case is selected for review by the case selection committee. This 90-day period includes the time required to translate information submitted by the person appealing, for Facebook to compile and submit supporting information, for the board members to deliberate their decision, and for Facebook to implement any decision.

8. What percentage of cases are overturned?

The Oversight Board selected its first cases in December 2020. As of April 30, 2021, the Oversight Board had issued decisions on eight cases, six of which overturned Facebook’s decisions, and two of which upheld Facebook’s decisions. One additional case became unavailable for review by the board as a result of user action.

9. Do individuals whose cases are taken up for review have the right to appear before the board?

For any case the board chooses to hear, the posting user is able to provide a statement on the case. The panel of members assigned to the case will receive information to help with their in-depth review, including information shared by the person who submitted the appeal as well as contextual information from Facebook, in compliance with applicable legal and privacy restrictions. The panel will deliberate the case and make a decision based on all the information provided by the person who submitted the appeal, by Facebook, and by any experts called upon to provide further context. Prior to publication, the posting user will also be able to choose if they want information that could identify them removed from the public case decision.

The Honorable Gus M. Bilirakis

The tech industry can often be a metrics-driven and data-driven world. Frequently, you're trying to improve metrics that measure engagement: video watch time for YouTube, the number of daily active users on Facebook, or the number of daily tweets on Twitter.

However, sometimes you can increase engagement by pushing both sides into their echo chambers. This echo chamber effect leads to less civil discourse and more political polarization, and it also allows misinformation to more easily to spread on both sides.

1. With respect to the problem of echo chambers, what metrics has your company developed which can measure this problem?

The question of whether social media has impacted political and social polarization has been the subject of swathes of serious academic research in recent years—the results of which are mixed, with many studies suggesting that social media is not the primary driver of polarization after all, and that evidence of the “echo chamber” effect is thin at best.

Research from Stanford last year looked in depth at trends in nine countries over 40 years and found that in some countries, polarization was on the rise before Facebook even existed, and in others it has been decreasing while internet and Facebook use increased. Other credible recent studies have found that polarization in the United States has increased the most among the demographic groups least likely to use the internet and social media, and data published in the EU suggests that levels of ideological polarization are similar whether you get your news from social media or elsewhere.

A Harvard study ahead of the 2020 US election found that election-related disinformation was primarily driven by elite and mass media, not least cable news, and suggested that social media played only a secondary role. And research from both Pew in 2019 and the Reuters Institute in 2017 showed that you're likely to encounter a more diverse set of opinions and ideas using social media than if you only engage with other types of media.

What evidence there is simply does not support the idea that social media, or the echo chamber it supposedly creates, is the unambiguous driver of polarization. One thing we do know is that political content is only a small fraction of the content people consume on Facebook—our own analysis suggests that in the US it is as little as 6%. Last year, Halloween had twice the increase in posting we saw on Election Day, and that's despite the fact that Facebook prompted people at the top of their News Feed to post about voting.

2. Mr. Zuckerberg, Facebook sets a limit at 5,000 friends, but not many average people realize this default setting. What are the meaningful distinctions and, perhaps needed default settings, for accounts that reach 500 versus 5 million people? In order words, what is you plan for reigning in super spreaders of harmful information?

When Pages, Groups, and websites repeatedly share content debunked by fact-checkers, we reduce the distribution of all of their posts in News Feed and remove them from the recommendations we show people. Pages may also lose their ability to monetize and advertise. For example, on 11 separate occasions, individuals identified by the Center for Countering

Digital Hate as spreading false or misleading information about COVID-19 and vaccines have had a presence on Facebook or Instagram permanently removed for violating our COVID-19 or vaccine misinformation policies.

We use a strike system, and all of our strikes are time-bound and situationally specific. When Pages and Groups have a certain number of violations in a certain period, additional enforcement measures are triggered.

Our strongest enforcement is reserved for content that has earned ratings of “False” and “Altered,” connoting the most outright false information. This enforcement can lead to the loss of monetization and advertising privileges. Content rated “Partly False” includes some factual inaccuracies, while the rating “Missing Context” is designed for content that may be misleading in the absence of additional context.

The Honorable Billy Long

- 1. Is it really your position that the FTC order from 2019, which was imposed as a result of your own violations of user privacy, can now be used to uphold a continued lack of transparency into your data?**

Transparency is one of the driving concepts behind our agreement with the FTC. We will have quarterly certifications to verify that our privacy controls are working. And where we find problems, we will make sure they're fixed. The process stops at the desk of our CEO, who signs a verification that we did what we said we would.

As a general matter, we strongly agree that the people who use Facebook should have a clear understanding of the kinds of data we collect and how we use that data. Privacy is at the core of everything we do, and our approach to privacy starts with our commitment to transparency and control—to helping people understand how their data is collected and used, and to giving them meaningful controls.

Our approach to transparency is threefold. First, wherever possible, we provide information about the data we collect and use and how people can control it in context and in our products. Research overwhelmingly demonstrates that in-product controls and education are the most meaningful to people and the most likely to be read and understood. Second, we provide information about how we collect and use data in our Data Policy, user agreements, and related educational materials. Third, we enable people to learn more about the specific data we have about them through interactive tools such as Download Your Information, which lets people download a file containing data that they may want to take to another service, and Access Your Information, a tool that lets people more easily access and manage their data on Facebook.

Our approach to control is based on the belief that people should be able to choose who can see what they share and how their data shapes their experience on Facebook. People can control the audience for their posts and the apps that can receive their data. They can control the people, Pages, Groups, and Events they connect to, and how they see content from those connections in their News Feed. They can provide feedback on every post they see on Facebook—feedback, for example, that they want to see less of a particular kind of post or fewer posts from a particular person or Page. They can see and delete the history of their activities on Facebook, and, if they no longer want to use Facebook, they can delete their account and the data associated with it.

In your lawsuit against BrandTotal, a startup marketing intelligence company, you allege that BrandTotal has violated your terms of service by collecting anonymized data from consenting users and you are therefore required to enforce the FTC order and revoke their access to your platform effectively killing the company. The FTC order was designed as a punishment to change your culture, not as a weapon for you to use to destroy competitive technologies.

- 2. Do you believe that the needs of your advertisers are more important than a person's right to consent to sharing their own data?**

We care deeply about privacy and are committed to our users' privacy and control over their data.

We filed a lawsuit against BrandTotal for using browser extensions designed to scrape data from Facebook, in violation of our Terms and Policies and state and federal law. The Court has already denied BrandTotal's request for a temporary restraining order and has rejected their competition claims and ruled they are unfounded.

Litigation is just one of the tools we use to combat scraping. We've also invested in technical teams and tools that monitor and detect suspicious activity and the use of unauthorized automation for scraping. This focus on scraping is part of our ongoing work to protect people's privacy.

It has been widely reported that during the fall, Facebook sent letters to certain researchers behind the NYU "Ad Observatory," alleging that the project violated provisions in its terms of service that prohibit bulk data collection from its site. Mr. Zuckerberg later suggested in testimony before the Senate Judiciary Committee that the NYU project "was scraping data in a way that we had agreed in our FTC consent decree around privacy that we would not allow" and accordingly Facebook had "to follow up on that" and "make sure that we take steps to stop that violation."

3. Provide the names of everyone who has received a cease and desist letter that mentions the FTC consent decree and provide copies of all such letters.

We recognize the importance of ads data to researchers and democracy, but we also have an obligation—including under our agreement with the FTC—to take steps to prevent unauthorized access to user data. In this case, NYU's browser extension works by scraping content from Facebook.

It is not our intention to impede transparency—far from it. We are committed to providing more transparency into the societal impact of Facebook's products while protecting people's privacy. Facebook's Ad Library and Ad Library API are used extensively by researchers to bring insights into political advertising. We're committed to providing ads transparency in ways that respect people's privacy, and we will continue to invest in new ways of doing that.

We have been working with interested researchers to identify ways that we can actively promote transparency and support their research goals while also protecting user privacy. For instance, the Facebook Open Research and Transparency (FORT) platform allows academic researchers to access data in a "virtual clean room," built with validated privacy and security protections. On February 1, we shared targeting information for US 2020 election ads with researchers for the first time. Around 1.65 million social issue, electoral, and political ads that ran in the US during the 90 days prior to Election Day (August 3 to November 3, 2020) were made available to researchers on the FORT platform. This is just one example of a step we're taking to help researchers access data they need to study Facebook's impact, and we're continuing to invest in resources that meet researchers' needs while putting privacy first.

4. **If Facebook has been in contact with the FTC about whether NYU or others are a “covered third party” under the consent decree, provide the details of such communications.**

While we are in regular contact with the FTC as a part of its oversight of our Consent Order, our practice is not to disclose the content of our confidential communications with FTC staff.

The Honorable Gary Palmer

There have been a lot of discussions about Russian and Chinese attempts to influence our elections and political discussions. I think we need to realize that these disinformation campaigns are not only aimed at us, but also our allies. For example, Radio Free Asia highlighted that China mobilized 300,000 cyber operatives to target our ally Taiwan's 2018 and 2020 elections. A 2016 Harvard study found that a group allied with the Chinese Communist Party (CCP) produced 488 million "fake" social media posts a year to distract other internet users from news and online discussions painting the Communist Party in a negative light.

1. What is your company doing outside of the U.S. to protect our allies like Taiwan from the nefarious activities of the CCP?

We continue to proactively investigate coordinated inauthentic behavior ("CIB") campaigns targeting public debate in the US and around the world. When we find CIB campaigns, we will investigate them, disrupt them, and make our findings and attribution public, as we've repeatedly committed to do. We'll also continue to share information about threats we disrupt on our platform with partner threat intelligence teams across industry and with key government partners. In August 2019, for example, we removed a network that originated in China and focused on Hong Kong. As another example, in September 2020, we removed a network that originated in China and focused primarily on the Philippines and Southeast Asia more broadly, and also on the United States.

2. What actions are you taking to keep individuals associated with the CCP from using your platform to attack the U.S. and our allies?

Please see the response to your previous question.

3. Do you pledge to work with our allies to push back against Russian and Chinese disinformation campaigns?

We work hard to reduce the spread of misinformation and disinformation on our platform. We've greatly expanded our efforts over the last few years to fight false news: we're getting better at enforcing against fake accounts and coordinated inauthentic behavior; we're using both technology and people to fight the rise in photo- and video-based misinformation; we've deployed new measures to help people spot false news and obtain more context about the stories they see in News Feed; and we've significantly grown our third-party fact-checking efforts.

Many media organizations, including state-sponsored media organizations, have a presence on Facebook. We apply labels to those state-controlled media outlets on Facebook and Instagram. Like everyone on Facebook, these organizations are required to comply with Facebook's policies, including complying with applicable laws. They must also comply with our policy requiring them to use their authentic names and identities. If we become aware that our policies are being violated, we will take action.

Our teams also continue to focus on finding and removing deceptive campaigns around the world—both foreign and domestic. For example, as outlined in our March 2021 Coordinated

Inauthentic Behavior Report (“CIB”), we removed fourteen networks of accounts, Pages, and Groups in March. We shared information about our findings with law enforcement, policymakers, and industry partners. At least nine out of the fourteen networks we investigated and took down—in Israel, Benin, Comoros, Georgia, and Mexico—targeted domestic audiences in their own countries. The other five—in Albania, Iran, Spain, Argentina, and Egypt—targeted primarily people outside of their countries. The vast majority of the networks we removed in March had limited following or were in the early stages of building their audiences when we removed them. We know these actors will continue to attempt to deceive and mislead people, including making particular viewpoints appear more widely supported or criticized than they are to sway public debate in their respective countries.

When it comes to influence operations, we’ve worked to identify new and emerging threats and remove coordinated inauthentic behavior across our platforms. In 2020 alone, we took down over 50 networks worldwide, many ahead of major democratic elections. As we’ve improved our ability to disrupt these operations, we’ve also built a deeper understanding of different threats and how best to counter them. Indeed, we continually adapt our defenses to make deceptive behaviors much more difficult and costly by building tools that make those tactics more difficult at scale. By continuing to develop smarter technologies, enhance our defenses, improve transparency, and build strong partnerships, we are making the constant improvements we need to stay ahead of our adversaries and to protect the integrity of our platforms.

The Honorable Michael C. Burgess, M.D.

According to Subsection (f)(3) of Section 230 of the Communications Decency Act, an “information content provider means any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service.”

Based on this almost anyone can be a content provider. One of my biggest concerns over the years has been the accessibility of Terms of Service and whether they have been applied equally to all users – users who must accept them in order to access your services.

- 1. On your platforms, if you limit distribution of user created content or include additional information to accompany a user’s content or post, are you then a content creator since you were “responsible...in part, for the...development of information”?**

We are, first and foremost, a technology company. Facebook does not create the content that our users publish on our platform. While we seek to be a platform for a broad range of ideas, we moderate content according to published Community Standards in order to keep the platform safe, reduce objectionable content, and to make sure people use the platform responsibly.

Section 230 of the Communications Decency Act provides that “[N]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.” Outside of certain specific exceptions, this means, among other things, that online platforms that host content created by others cannot be held liable for that content. Likewise, Section 230 provides that a platform cannot be held liable for a decision to remove or restrict access to content created by others.

With respect to our work around misinformation more generally, people often tell us they don’t want to see misinformation. We work with over 80 independent third-party fact-checkers who are certified through the non-partisan International Fact-Checking Network (IFCN) to help identify and review false news. If content is deemed by a fact-checker to be False, Altered, or Partly False, according to our public definitions, its distribution will be reduced, and it will appear lower in News Feed. We also implement an overlaid warning screen on top of fact-checked content.

- 2. Should Congress be considering an update to the definition of an “information content provider”?**

Section 230 of the Communications Decency Act has been essential to protecting free expression and innovation on the internet. Section 230 allows us to do the work we need to do to help keep people safe by enforcing our Community Standards.

That said, as we did in the case of SESTA/FOSTA, we would welcome the opportunity to work with Congress on thoughtful proposals to modify Section 230. Identifying a way forward is challenging given the chorus of people arguing—sometimes for contradictory reasons—that the law is doing more harm than good. Facebook stands ready to be a productive partner in the discussion about Section 230 reform.

Section 230 has been hailed as the 26 words that created the Internet and vital to the success of small and medium sized online businesses. Your platform has benefited greatly from Section 230 liability protections.

3. What would your service look like without Section 230?

At a high level, Section 230 does two things. First, it encourages free expression. Without Section 230, platforms could potentially be held liable for everything people say. Without this protection, platforms would likely remove more content to avoid legal risk and would be less likely to invest in technologies that enable people to express themselves in new ways. Second, it allows platforms to remove harmful content. Without Section 230, platforms could face liability for doing even basic moderation, such as removing bullying and harassment that impact the safety and security of their communities. Repealing Section 230 entirely would also likely substantially increase many companies' costs associated with legal challenges and content moderation.

4. What about when you were first starting out?

Section 230 made it possible for a broad range of online services to be built and ensured important values like free expression and openness were part of how platforms operate. As the internet keeps growing and evolving, the core principles of Section 230 will continue to be crucial for innovation—for small platforms that don't have the same capabilities when it comes to content moderation, for large ones that host billions of pieces of content across the globe, and for the American tech sector as a whole if we are going to maintain our edge in innovation. But that doesn't mean it shouldn't be updated to reflect the way the internet has changed in the last 25 years—and that's why we support thoughtful reform to make sure the law is working as intended. We've been clear that the recommendations we've made for Section 230 reform should only apply to larger platforms. It's really critical that a small platform—the next student in a dorm room or garage—has a relatively low regulatory burden in order to be able to innovate and then get to the scale where they can afford to put those kinds of systems in place.

Mr. Zuckerberg, to use Facebook, users must accept your Terms of Service that include Facebook's Community Standards. Your Community Standards state that they "apply to everyone, all around the world, and to all types of content." However, given the billions of interactions on Facebook each day, we know that not every post that violates your Community Standards is removed or addressed.

5. Are you taking any steps to increase monitoring and enforcement of your Community Standards to achieve equal application to all users?

Yes, this is an area in which we have invested heavily. In the last few years we have tripled—to more than 35,000—the number of people working on safety and security. This includes our growing team of more than 15,000 content reviewers—a mix of full-time employees, contractors, and companies we partner with. Our reviewers come from many backgrounds, reflect the diversity of our community, and bring a wide array of professional experiences, from veterans to former public sector workers.

In addition to our content review teams, to enforce our Community Standards, we have introduced tools that allow us to proactively detect and remove violating content using advances in technology, including artificial intelligence, machine learning, and computer vision. We do this by analyzing specific examples of bad content that have been reported and removed to identify patterns of behavior. Those patterns can be used to teach our software to proactively identify similar content. These advances in technology mean that we can now remove bad content more quickly, identify and review more potentially harmful content, and increase the capacity of our review team. To ensure the accuracy of these technologies, we constantly test and analyze our systems, technology, and AI to ensure accuracy. All content goes through some degree of automated review, and we use human reviewers to check some content that has been flagged by that automated review or reported by people that use Facebook. We also use human reviewers to perform reviews of content that was not flagged or reported by people to check the accuracy and efficiency of our automated review systems. The percentage of content that is reviewed by a human varies widely depending on the type and context of the content, and we don't target a specific percentage across all content on Facebook.

Our Community Standards Enforcement Report shares metrics on how Facebook is performing in preventing and removing content that violates our Community Standards. The report specifies how much content we took action on during the specified period, as well as how much of it we found before users reported it to us.

Mr. Zuckerberg, my understanding is that automated systems do the initial scanning of user posts to determine if they violate the Community Standards. Some posts are then flagged and reviewed by humans. When a human reviewer makes a decision on allowing a user's content to remain on Facebook, the decision is inherently subjective, even if based on the reviewer's understanding of the User Agreement.

6. Do you believe these decisions render Facebook a publisher rather than a platform in the context of Section 230?

Please see the response to your Question 1.

If enforcement of Terms of Service is subjective – due to a human reviewer – at a high enough frequency that application of the Terms is inherently unequal across all users, then I would argue that the Terms need adjusting to be universally objective.

7. Would universally objective Terms of Service result in virtually no Community Standards?

Our Community Standards are global, and all reviewers use the same guidelines when making decisions. And our reviewers are trained to look for violations and enforce our policies as consistently and as objectively as possible. They undergo extensive training when they join and thereafter are regularly trained and tested with specific examples on how to uphold the Community Standards and take the correct action on a piece of content. This training includes when policies are clarified or as they evolve.

We seek to write actionable policies that clearly distinguish between violating and non-violating content, and we seek to make the decision-making process for reviewers as objective as

possible. There are quality control mechanisms as well as management on site to help or seek guidance from if needed. When a reviewer isn't clear on the action to take based on the Community Standards, they can pass the content decision to another team for review. We also audit the accuracy of reviewer decisions on an ongoing basis to coach them and follow up on improving where errors are being made. When we're made aware of incorrect content removals, we review them with our Community Operations team so as to prevent similar mistakes in the future. We have also introduced the right to appeal our decisions on individual posts so users can ask for a second opinion when they think we've made a mistake.

Facebook's Terms and Policies are available here: <https://www.facebook.com/policies>. Facebook's Community Standards are available at <https://www.facebook.com/communitystandards/>. Both are global and apply to all accounts on Facebook.

Additionally, if someone believes that we have gotten a content moderation decision wrong, that person can generally appeal or is given the option to disagree with our decision. In most cases, we also provide users with the reason for our decision and the specific policy that was violated. In some cases, we then re-review our decisions on those individual pieces of content.

In order to request re-review of a content decision we made, users are often given the option to "Request Review" or to provide feedback by stating they "Disagree with Decision." We try to make the opportunity to request this review or give this feedback clear, either via a notification or interstitial, but we are always working to improve.

Transparency in our appeals process is important, so we now include in our Community Standards Enforcement Report how much content people appealed and how much content was restored upon appeal. Gathering and publishing those statistics keeps us accountable to the broader community and enables us to continue improving our content moderation. For more information, see <https://transparency.facebook.com/community-standards-enforcement>.