

OUR POSITION ON POLARIZATION AND ELECTIONS

You will have seen the series of articles about us published in the Wall Street Journal in recent days, and the public interest it has provoked. This Sunday night, the ex-employee who leaked internal company material to the Journal will appear in a segment on 60 Minutes on CBS. We understand the piece is likely to assert that we contribute to polarization in the United States, and suggest that the extraordinary steps we took for the 2020 elections were relaxed too soon and contributed to the horrific events of January 6th in the Capitol.

I know some of you – especially those of you in the US – are going to get questions from friends and family about these things so I wanted to take a moment as we head into the weekend to provide what I hope is some useful context on our work in these crucial areas.

Facebook and Polarization

People are understandably anxious about the divisions in society and looking for answers and ways to fix the problems. Social media has had a big impact on society in recent years, and Facebook is often a place where much of this debate plays out. So it's natural for people to ask whether it is part of the problem. But the idea that Facebook is the chief cause of polarization isn't supported by the facts – as Chris and Pratiti set out [in their note on the issue](#) earlier this year.

The rise of polarization has been the subject of swathes of serious academic research in recent years. In truth, there isn't a great deal of consensus. But what evidence there is simply does not support the idea that Facebook, or social media more generally, is the primary cause of polarization.

The increase in political polarization in the US pre-dates social media by several decades. If it were true that Facebook is the chief cause of polarization, we would expect to see it going up wherever Facebook is popular. It isn't. In fact, polarization has gone down in a number of countries with high social media use at the same time that it has risen in the US.

Specifically, we expect the reporting to suggest that a change to Facebook's News Feed ranking algorithm was responsible for elevating polarizing content on the platform. In January 2018, we made ranking changes to promote Meaningful Social Interactions (MSI) – so that you would see more content from friends, family and groups you are part of in your News Feed. [This change was heavily driven by internal and external research](#) that showed that meaningful engagement with friends and family on our platform was better for people's wellbeing, and we further refined and improved it over time as we do with all ranking metrics. Of course, everyone has a rogue uncle or an old school classmate who holds strong or extreme views we disagree with – that's life – and the change meant you are more likely to come across their posts too. Even so, we've developed industry-leading tools to remove hateful content and reduce the distribution of problematic content. As a result, the prevalence of hate speech on our platform is now down to about 0.05%.

But the simple fact remains that changes to algorithmic ranking systems on one social media platform cannot explain wider societal polarization. Indeed, polarizing content and misinformation are also present on platforms that have no algorithmic ranking whatsoever, including private messaging apps like iMessage and WhatsApp.

Elections and Democracy

There's perhaps no other topic that we've been more vocal about as a company than on our work to dramatically change the way we approach elections. Starting in 2017, we began building

new defenses, bringing in new expertise, and strengthening our policies to prevent interference. Today, we have more than 40,000 people across the company working on safety and security. Since 2017, we have disrupted and removed more than 150 covert influence operations, including ahead of major democratic elections. In 2020 alone, we removed more than 5 billion fake accounts — identifying almost all of them before anyone flagged them to us. And, from March to Election Day, we removed more than 265,000 pieces of Facebook and Instagram content in the US for violating our voter interference policies.

Given the extraordinary circumstances of holding a contentious election in a pandemic, we implemented so called “break glass” measures – and [spoke publicly about them](#) – before and after Election Day to respond to specific and unusual signals we were seeing on our platform and to keep potentially violating content from spreading before our content reviewers could assess it against our policies.

These measures were not without trade-offs – they’re blunt instruments designed to deal with specific crisis scenarios. It’s like shutting down an entire town’s roads and highways in response to a temporary threat that may be lurking somewhere in a particular neighborhood. In implementing them, we know we impacted significant amounts of content that did not violate our rules to prioritize people’s safety during a period of extreme uncertainty. For example, we limited the distribution of live videos that our systems predicted may relate to the election. That was an extreme step that helped prevent potentially violating content from going viral, but it also impacted a lot of entirely normal and reasonable content, including some that had nothing to do with the election. We wouldn’t take this kind of crude, catch-all measure in normal circumstances, but these weren’t normal circumstances.

We only rolled back these emergency measures – based on careful data-driven analysis – when we saw a return to more normal conditions. We left some of them on for a longer period of time through February this year and others, like not recommending civic, political or new Groups, we have decided to retain permanently.

Fighting Hate Groups and other Dangerous Organizations

I want to be absolutely clear: we work to limit, not expand hate speech, and we have clear policies prohibiting content that incites violence. We do not profit from polarization, in fact, just the opposite. We do not allow dangerous organizations, including militarized social movements or violence-inducing conspiracy networks, to organize on our platforms. And we remove content that praises or supports hate groups, terrorist organizations and criminal groups.

We’ve been more aggressive than any other internet company in combating harmful content, including content that sought to delegitimize the election. But our work to crack down on these hate groups was years in the making. We took down tens of thousands of QAnon pages, groups and accounts from our apps, removed the original [#StopTheSteal](#) Group, and removed references to Stop the Steal in the run up to the inauguration. In 2020 alone, we removed more than 30 million pieces of content violating our policies regarding terrorism and more than 19 million pieces of content violating our policies around organized hate in 2020. We designated the Proud Boys as a hate organization in 2018 and we continue to remove praise, support, and representation of them. Between August last year and January 12 this year, we identified nearly 900 militia organizations under our Dangerous Organizations and Individuals policy and removed thousands of Pages, groups, events, Facebook profiles and Instagram accounts associated with these groups.

This work will never be complete. There will always be new threats and new problems to address, in the US and around the world. That's why we remain vigilant and alert – and will always have to.

That is also why the suggestion that is sometimes made that the violent insurrection on January 6 would not have occurred if it was not for social media is so misleading. To be clear, the responsibility for those events rests squarely with the perpetrators of the violence, and those in politics and elsewhere who actively encouraged them. Mature democracies in which social media use is widespread hold elections all the time – for instance Germany's election last week – without the disfiguring presence of violence. We actively share with Law Enforcement material that we can find on our services related to these traumatic events. But reducing the complex reasons for polarization in America – or the insurrection specifically – to a technological explanation is woefully simplistic.

We will continue to face scrutiny – some of it fair and some of it unfair. We'll continue to be asked difficult questions. And many people will continue to be skeptical of our motives. That's what comes with being part of a company that has a significant impact in the world. We need to be humble enough to accept criticism when it is fair, and to make changes where they are justified. We aren't perfect and we don't have all the answers. That's why we do the sort of research that has been the subject of these stories in the first place. And we'll keep looking for ways to respond to the feedback we hear from our users, including testing ways to make sure political content doesn't take over their News Feeds.

But we should also continue to hold our heads up high. You and your teams do incredible work. Our tools and products have a hugely positive impact on the world and in people's lives. And you have every reason to be proud of that work.