

The Washington Post

Facebook's new 'Supreme Court' knocks content rules as vague and incomplete in first rulings

Dwoskin and Craig Timberg

Jan. 28, 2021

Facebook's new Oversight Board found company content moderation policies vague and poorly communicated in its first [set of decisions released Thursday](#), overruling the company's actions in four of the five cases it decided in its initial round of cases.

The actions covered a range of issues that have vexed social media companies — alleged hate speech, [coronavirus](#) misinformation and references to dangerous organizations and people — and included one case in which an automated detection system apparently overreacted to an image of an uncovered female nipple in a breast cancer awareness campaign.

Taken together, the rulings suggest the oversight board is going to demand greater clarity and transparency from Facebook in the tiny sliver of cases it chooses to review. The board is also weighing Facebook's ban of President Donald Trump following the Jan. 6 riot at the U.S. Capitol, though a decision in that case is not likely for months. The five cases decided Thursday all date to October or November of last year.

"We often found that the community standards as written are incomplete," board member Sudhir Krishnaswamy, vice-chancellor of the National Law School of India University, said in an interview with The Washington Post.

He added that the cultural and linguistic contexts of users around the world can make moderation difficult and that the case-by-case nature of Facebook's policy development may have hindered the creation of clear, coherent policies over time. He said the Oversight Board's demand for better explanation and increased rigor is likely to help spur better policies overall and may improve the approach of other technology companies.

"I suspect [such a careful review] has not happened before with any of the companies," Krishnaswamy said. "I suspect this is a wide problem with social media across the Internet."

The board issued nine policy recommendations in addition to the rulings. Facebook has seven days to restore the removed content, but the company said Thursday morning it already had acted to restore the content in all four cases in which its actions were overruled. It also will look to see if similar content from other users should be restored and will consider the policy recommendations from the board.

"We believe that the board included some important suggestions that we will take to heart," Monika Bickert, vice president of content policy, said in a [Facebook blog post](#). "Their recommendations will have a lasting impact on how we structure our policies."

The six cases were selected from 150,000 submissions from across four continents of different instances where users believed content was unfairly removed.

“None of these cases had easy answers, and deliberations revealed the enormous complexity of the issues involved,” the board said in a [blog post](#) Thursday morning summarizing its actions.

The board — which was launched last year and is funded by Facebook — is intended to function as a “Supreme Court” where the toughest decisions about free expression online can be decided and is considered a potential alternative to the regulation of the social media industry that is being considered by governments all over the world, including the United States. It’s composed of 20 members, including a former prime minister, a Nobel laureate, as well as journalists and legal experts from 16 nations.

The board has the power to make Facebook change its content decisions on specific issues, but has been criticized because it cannot directly change Facebook’s policies going forward. The board can issue recommendations for policy changes that could impact billions of users in the future, but Facebook is not required to implement them.

In an interview, board member John Samples, a vice president at the libertarian-leaning think tank Cato Institute, said the decisions announced Thursday show that “the board is not willing to let Facebook off the hook.”

He said that he hoped the board would demonstrate that a model for fair governance of social media could exist outside government regulation, and that he was drawn to the idea of shaping a system for online expression that was still evolving. “This is a 10-year development toward what we hope will be the right answer.”

The idea for an external oversight board was first floated by Facebook CEO Mark Zuckerberg in 2018. He said at the time that he did not believe it made sense for critical content decisions to be concentrated in the hands of one company. Zuckerberg has promised to abide by the board’s rulings, which the company called “binding decisions” in its blog post Thursday, but it is under no legal obligation to do so.

Zuckerberg has said that he supports government regulation of the social media industry, which the Biden administration and other government are considering as they wrestle with companies that have enormous power to control the free expression of billions of people.

Officials and policymakers all over the world who are seeking to design new frameworks for regulating the social media industry are watching the board closely. If it’s judged a success, it may lessen the calls for regulation. If it fails, the failure may hasten demands to create more-stringent legal guardrails for content moderation in many countries.

Facebook and other social media companies over the last year have been more aggressive than ever before about policing speech and have enacted first-time policies banning misinformation about the coronavirus and about the U.S. presidential election. These unprecedented efforts, while largely unsuccessful in preventing the spread of misinformation, have made questions about the role of private companies in policing content even more pressing.

The Oversight Board operates through five-person panels, one member of which has to be from the region where a particular case originates. The board and its staff select the cases, not Facebook. The panels then review comments on each case, consult experts and make

recommendations to the full board, which makes final decisions through majority vote. Deliberations so far have been online because of pandemic-related restrictions on travel and meeting in person.

In one of the five cases made public Thursday, the board upheld a Facebook decision to remove a post referring to Azerbaijanis by what the board agreed was a “dehumanizing slur attacking national origin.” It said the action correctly applied Facebook policies to protect the safety and dignity of people even if such actions undermine a user’s “voice.”

But the board found problems in four other cases, including the one removing images of nipples in the breast cancer awareness campaign in Brazil. An automated system took this action on Instagram, which Facebook owns, and the company already had reversed it. The company already counts “breast cancer awareness” as an exception to its policy prohibiting nudity — but the board continued to review the case to make the point that Facebook’s automated systems are problematic.

A case on hate speech dealt with a post from a user in Myanmar suggesting that there is “something wrong with Muslims (or Muslim men) psychologically or with their mindset.” But the board questioned the accuracy of Facebook’s translation of the post and ruled that its full context “did not advocate hatred or intentionally incite any form of imminent harm.”

The board similarly found that user who incorrectly quoted Joseph Goebbels, a Nazi propaganda chief, did not in fact violate Facebook’s policy on dangerous individuals and organizations because the quote did not support Nazi ideology or actions. The board also called on Facebook to make clearer to users what statements would violate this policy and to provide examples.

In another case, the board found Facebook incorrectly removed a piece of content in which a user criticized the French government’s policies on the coronavirus. In the banned post, the user complained that the French government’s refusal to authorize the anti-malaria drugs hydroxychloroquine and azithromycin was problematic because such drugs were “being used elsewhere to save lives.”

Facebook removed the post on the grounds that encouraging people to take an unproven drug for covid-19 could cause imminent harm to people.

The board overturned that determination, arguing that Facebook failed to define or demonstrate how encouraging people to take a drug that can’t be obtained without a prescription in France could cause “imminent” harm. The board also said that Facebook had failed to create clear rules of the road for health misinformation, noting that it was not logical for the social network to equate every single piece of misinformation about covid-19 treatments or cures as “necessarily rising to the level of imminent harm.” Facebook’s own policies say that additional context is needed before the company will remove content on such grounds, the board noted in its decision.

The board also recommended that Facebook create a more nuanced system of enforcement to tackle coronavirus- and health-related misinformation, a recommendation that Facebook can adopt voluntarily if it chooses.