



## Public Comment Appendix for

2021-004-FB-UA

Case number

Case description

In January 2021, User A left a comment summarising their first-hand experience during the recent protests in support of the Russian opposition leader Alexei Navalny. In the comment, User A called User B – who criticised the protesters – a "common cowardly bot" [банальный трусливый бот].

On 24 January, User B commented on a post consisting of several pictures, a video and text of the protests in support of Alexei Navalny held in Saint Petersburg and across Russia on 23 January. User B claimed that they did not know what happened in Saint Petersburg, but that protesters in Moscow were all school children, mentally "slow" and were "used". They added that the protesters were not the voice of the people but a "theatre show". Other users began challenging this critical account of the protesters in the comments thread.

User A challenged User B's assertion. User A identified as elderly and claimed to have participated in the Saint Petersburg protests with their colleagues and adult children. User A stated that there were so many protesters that they lost their companions in the crowd. They claimed to have witnessed elderly and disabled protesters. They expressed pride in the youth that participated in the protests and dismissed the claim that anyone manipulated them. In their final words, User A called User B a "common cowardly bot" ["банальный трусливый бот"].

User B was the only person who reported the comment. Facebook took it down under its [Bullying and Harassment Community Standard](#), which provides for the removal of content that is meant to degrade or shame private individuals. In certain instances, Facebook requires self-reporting by the person who has been targeted by the bullying or harassment.

As part of their appeal, User A explained to the Board that they had shared first-hand experience from the protests and that they were responding to blatantly false information. User A added that they believed that User B was "a bot that works on order without being an eyewitness and participant in the events". User A further indicated that their opinion about this user had not changed, that the term "bot" was not a "dirty word" and that they did not endorse terrorism or any illegal action.

The Board would appreciate public comments that address:

- Whether Facebook's decision complied with its Community Standard on Bullying and Harassment.

- Whether Facebook's decision and Community Standard on Bullying and Harassment complied with its values and its human rights responsibilities, considering both the importance of free expression and the need to ensure scalable responses to harm caused by bullying and harassment.
- Contextual information regarding the common use or acceptance of terms equivalent to "банальный трусливый бот" ("common cowardly bot") when discussing emotive topics on social media, specifically in the Russian language.
- Any effects of Facebook's enforcement of its Community Standards on dissenting or minority political viewpoints in contexts where governments routinely restrict critical expression.
- Research on dis- and misinformation campaigns against participants in the January protests in Russia, how it is disseminated on social media, including evidence of any coordinated inauthentic behaviour and the actors involved.



Public Comment Appendix for  
2021-004-FB-UA

Case number

The Oversight Board is committed to bringing diverse perspectives from third parties into the case review process. To that end, the Oversight Board has established a public comment process.

Public comments respond to case descriptions based on the information provided to the Board by users and Facebook as part of the appeals process. These case descriptions are posted before panels begin deliberation to provide time for public comment. As such, case descriptions reflect neither the Board's assessment of a case, nor the full array of policy issues that a panel might consider to be implicated by each case.

To protect the privacy and security of commenters, comments are only viewed by the Oversight Board and as detailed in the [Operational Privacy Notice](#). All commenters included in this appendix gave consent to the Oversight Board to publish their comments. For commenters who did not consent to attribute their comments publicly, names have been redacted. To withdraw your comment, please email [contact@osbadmin.com](mailto:contact@osbadmin.com).

To reflect the wide range of views on cases, the Oversight Board has included all comments received except those clearly irrelevant, abusive or disrespectful of the human and fundamental rights of any person or group of persons and therefore violating the [Terms for Public Comment](#). Inclusion of a comment in this appendix is not an endorsement by the Oversight Board of the views expressed in the comment. The Oversight Board is committed to transparency and this appendix is meant to accurately reflect the input we received.



Public Comment Appendix for  
2021-004-FB-UA

Case number

23

Number of Comments

Regional Breakdown

0	1	8	1
Asia Pacific & Oceania	Central & South Asia	Europe	Latin America & Caribbean
0	0	13	
Middle East and North Africa	Sub-Saharan Africa	United States & Canada	

**2021-004-FB-UA**

Case number

**PC-09996**

Public comment number

**United States and Canada**

Region

**Leslie**

Commenter's first name

**Comar Tudela**

Commenter's last name

**English**

Commenter's preferred language

**DID NOT PROVIDE**

Organization

**No**

Response on behalf of organization

-----

Short summary provided by the commenter

**I support democracy and Navaly in Russia and think that posts supporting both including protests should not be removed by Facebook because Truth should never be censored only lies and disinformation**

Full Comment

**I support democracy both in the United States and in Russia and support Navaly and what he is trying to do at all personal costs to him and his family. Putin's dominant tyrannical rule has made him rich on the backs of his people and his power is far to great for far too long. Truth to power and only lies and disinformation should be deleted by Facebook Stop censoring the truth and promoting lies**

Link to Attachment

**No Attachment**

**2021-004-FB-UA**

Case number

**PC-09997**

Public comment number

**United States and Canada**

Region

**Sandra**

Commenter's first name

**Yukman**

Commenter's last name

**English**

Commenter's preferred language

**DID NOT PROVIDE**

Organization

**No**

Response on behalf of organization

-----

Short summary provided by the commenter

While I believe this is a case of misinformation and the protestors were not manipulated children, ( had to research and find that out on my own), it isn't bullying to call someone a name like a bot. It's sad but doesn't rise to the level of removal the post. My wish is that Facebook, like Twitter, would take some of its revenue and fact check things like this that are highly politically explosive and have the potential to destroy many many lives.

Full Comment

There is plenty of information available to show that user B was spreading misinformation. FB should simply point this out by linking to a Yisrael source so that potentially millions are not led at astray by these comments.

Link to Attachment

**No Attachment**

**2021-004-FB-UA**

Case number

**PC-10000**

Public comment number

**United States and Canada**

Region

**Dr. Brett**

Commenter's first name

**Prince**

Commenter's last name

**English**

Commenter's preferred language

**DID NOT PROVIDE**

Organization

**No**

Response on behalf of organization

-----

Short summary provided by the commenter

**It should not have been taken down.**

Full Comment

**There is nothing violative of community standards. This political and personal difference of opinion/perception does not rise to that level. It should not have been taken down. Free and open speech should be honored and respected.**

Link to Attachment

**No Attachment**

**2021-004-FB-UA**

Case number

**PC-10003**

Public comment number

**United States and Canada**

Region

**Holly**

Commenter's first name

**Fultz**

Commenter's last name

**English**

Commenter's preferred language

**DID NOT PROVIDE**

Organization

**No**

Response on behalf of organization

-----

Short summary provided by the commenter

If User A was truthful in stating "he" was in fact at the protest, "he" has a right to respond stating such as fact and discredit statements insinuating a fact which is not supported by their actual experience. However, name calling "common cowardly bot" is a form of bullying therefore against facebook antibullying policy.

Full Comment

If User A was truthful in stating "he" was in fact at the protest, "he" has a right to respond stating such as fact and discredit statements insinuating a fact which is not supported by their actual experience (User B). However, name calling "common cowardly bot" is a form of bullying therefore

Link to Attachment

**No Attachment**



2021-004-FB-UA

Case number

PC-10005

Public comment number

United States and Canada

Region

Robert

Commenter's first name

Kriegar

Commenter's last name

English

Commenter's preferred language

DID NOT PROVIDE

Organization

No

Response on behalf of organization

-----

Short summary provided by the commenter

Another case where Facebook penalized the victim rather than the attacker. A common issue.

Full Comment

"User B claimed that they did not know what happened in Saint Petersburg, but that protesters in Moscow were all school children, mentally "slow", and were "used." They added that the protesters were not the voice of the people but a "theatre show." " What is clear is that User B is guilty of "violating community guidelines on bullying and harassment", not User A. User B has chosen to use derogatory terms and demean users who disagree with him, based upon terms meant to demean and deride people based upon mental handicap, intellect, or diminished capacity. In fact, User B seems challenged in avoiding the direct word he wants to convey by inference, "retard", an easily identified slur. User A, however, seeks to identify User B as a known class of poster, and uses a valid characterization of that opinion. The term "not" is neither demeaning or deriding. It is simply a known form of poster. As is a "troll", or "trolling". User B, as is a known habit of trolls and bots, wishes to project his conduct upon User A, who oddly enough is immediately penalized by Facebook and who is also accused of being guilty upon review, incorrectly. Without the Oversight Boards involvement, User A would have no further recourse or ability to appeal, and could find himself subject to penalty with no recourse. Including having his account suspended for 30 day stretches, effectively silencing him and stripping him of his right to freely and reasonably express himself. Another known habit Facebook employs to unfairly censure any but far right wing users. Meanwhile, User B enjoys the ability to continue trolling, and bullying others using known terms of hate speech to demean others on the basis of mental disease or defect. Why User B is not censored and suspended for his bullying is beyond comprehension, and Facebook, as usual, has not taken User B's comments into account, nor has Facebook taken User A's completely acceptable assertion in

context. The effect of this is that it nullifies, diminishes, and contravenes Facebook's STATED policies on bullying and harassment, and in fact, reinforces the actions of the bully, and magnifies the victimization of the victims. This ends up in negatively impacting ALL valid Facebook users in a negative way, and stultified the free exchange of ideas and self-expression. The course of this activity is so widespread and so well known that it also negatively impacts the Facebook brand and it's value to real users and advertisers. The real world value of Facebook to actual living person's is also damaged and lessened. Therefore, what could be a powerful tool for the population now becomes a tool for entities other than the populace to control and silence the free exchange of ideas among the populace, and control political expression in a false way.

[Link to Attachment](#)

**No Attachment**

**2021-004-FB-UA**

Case number

**PC-10006**

Public comment number

**Europe**

Region

**Kimberly**

Commenter's first name

**Meadows**

Commenter's last name

**English**

Commenter's preferred language

**DID NOT PROVIDE**

Organization

**No**

Response on behalf of organization

-----

Short summary provided by the commenter

**I think that a ban for this comment is too harsh especially as only one person reported the comment. I have seen much worse comments that have been deemed ok by community standards.**

Full Comment

**So in summary, I don't believe the comment to be harmful to anybody, or should be classified as bullying.**

Link to Attachment

**No Attachment**

2021-004-FB-UA

Case number

PC-10008

Public comment number

United States and Canada

Region

Lori

Commenter's first name

Chinitz

Commenter's last name

English

Commenter's preferred language

OGreat One

Organization

No

Response on behalf of organization

-----

Short summary provided by the commenter

I do not believe that the comment “common cowardly bot” [банальный трусливый бот] comes anywhere near the standard of online bullying and should not have been taken down.

Full Comment

The comment that User A made (“common cowardly bot” [банальный трусливый бот]) does not rise to the level of bullying. It is a relatively mild insult. User A was surmising that User B was--if not a bot--a Russian government asset--and leveled this insult at him. The fact that the power, in this situation, lies in the Russian government's camp, should make the Oversight Board very cautious about banning the speech of the opposition--especially such a pedestrian insult. I think that most people would agree that the Russian side is more of a Goliath than a David in this case. I am not a fluent Russian speaker but have basic conversational comprehension. I would wager that most insults available to Russian speakers would pack more of a punch than "banal cowardly bot." The fact that User B insinuated that the protesters were (en masse) mentally defective should have set off more alarm bells than the comment of a protester that his critic might be a bot or a toady. Please reserve your bans for veritable hate speech.

Link to Attachment

No Attachment

2021-004-FB-UA

Case number

PC-10011

Public comment number

United States and Canada

Region

Withheld

Commenter's first name

Withheld

Commenter's last name

English

Commenter's preferred language

Withheld

Organization

No

Response on behalf of organization

-----

Short summary provided by the commenter

This case has to do with Facebook's community standards on bullying and harassment. I feel that this system is broken, as the automatic system flags innocent comments all the time, unfairly banning people from use of Facebook without any recourse. I am on a personal mission to bring this issue to Facebook's attention, but it is hard to get the ear of a company that does not like to listen.

Full Comment

"Common cowardly bot" is no more offensive than anything I've been 'subjected' to. While this case deals with a protest in Saint Petersburg in support of Alexei Navalny, it also has parallels with a lot of contentious political discourse throughout the world. Additionally, I've come across bots on Facebook and other social media, doing a variety of things. Facebook's system of automatic banning already produces far too many false positives, such as when two friends call each other "silly goose" or "pinecone". These are not offensive terms, these are not cases of bullying, yet Facebook bans people for days over using these terms without any recourse. When the user asks for a review, it is evident that it is not reviewed by a person at all. So if Facebook were to start banning for calling out bots, I don't think the system works well enough to differentiate between calling another person a "cowardly bot" and tagging someone in a comment letting them know that they appear to have been hacked by a bot (this is also a common occurrence on Facebook). It would generate even more false positives. I strongly feel that in ruling on this case, Facebook's community standards should be overhauled so that only actual cases of bullying result in a suspension. The automatic system should be taken offline entirely, as it produces too many false positives, until it can be reprogrammed so that only the most serious of comments gets flagged. There MUST be a way to provide customer support and recourse for users banned for a few days to 30 days, especially when the reason for the ban is a false positive. The automatic system also has no way of

differentiating dialects, so something that may be completely innocent in one country might be extremely offensive in another. The word "negro" is used for the color "black" in Puerto Rican Spanish, and I heard of someone in a Puerto Rican bird watching group getting a ban for talking about "finche negro" as in the bird "black finch". By banning that person, Facebook was acting in a racist and insensitive manner. Facebook was the one engaging in bullying and harassment, not the person talking about a bird. I think the automatic system should be reserved for clear and flagrant violations (a racial slur that is considered as such universally, such as the N-word), and that additional violations could just go back to relying on user reports. Perhaps the automatic system should also be tooled to detect the language, so false positives don't occur as the result of a foreign language having a coincidental word to a slur (Norwegian has a word 'fag' that means 'subject')

Although personally, I think all things should be left up to users to report, and the automatic system taken down entirely. I once got a ban for posting a Youtube link to Soul Asylum's Runaway Train. My crime? I posted the link to just the song, which showed the album art, and that had my post flagged as "nudity". Yet weeks before, Facebook had been banning people for posting Houses of the Holy by Led Zeppelin, but fixed that obvious error.

[Link to Attachment](#)

**No Attachment**

**2021-004-FB-UA**

Case number

**PC-10012**

Public comment number

**United States and Canada**

Region

**Withheld**

Commenter's first name

**Withheld**

Commenter's last name

**English**

Commenter's preferred language

**Withheld**

Organization

**No**

Response on behalf of organization

-----

Short summary provided by the commenter

Your bullying rules should be the same as the US supreme court ruled on in the 1970's as required in order to suppress free speech under the 1st amend. that is the suppressed speech must be a threat of IMMEDIATE harm to a specific person. Perhaps, the term immediate could be loosened with social media, but there must be discernable, physical harm to a person threatened in order to be deleted in a civil, free democracy. otherwise, we live in a totalitarian society

Full Comment

**my full comment is above**

Link to Attachment

**No Attachment**

**2021-004-FB-UA**

Case number

**PC-10017**

Public comment number

**United States and Canada**

Region

**Withheld**

Commenter's first name

**Withheld**

Commenter's last name

**English**

Commenter's preferred language

**Withheld**

Organization

**No**

Response on behalf of organization

-----

Short summary provided by the commenter

**Was not bullying or harrasment**

Full Comment

The comment did not appear to be bullying or harassment, which more indicate unreasonable abusive or continued force. Calling someone making a comment a bot may be making a true statement, unless Facebook wants to certify otherwise. The coward statement may also be true, if no true name was given; or, it might be a way to goad someone into giving a name in order to retaliate later. If the latter is Facebook's concern, just ban the word coward when used as at a person in the conversation. Software can probably do that; but, it seems a largely harmless term to make such effort worth little. The typical person without a complex should be able to handle such language.

Link to Attachment

**No Attachment**



2021-004-FB-UA

Case number

PC-10019

Public comment number

United States and Canada

Region

Withheld

Commenter's first name

Withheld

Commenter's last name

English

Commenter's preferred language

Withheld

Organization

No

Response on behalf of organization

-----

Short summary provided by the commenter

That User A was not bullying in their reported response, as User B appears to be falsely attempting to smear a valid human rights movement. The current Russian Gov't has often used bots to spread disinformation, so the comment is reasonable. By taking down the post by User A, Facebook is harming political dissent within Russia.

Full Comment

The comments by User B are in my opinion consistent with the style I have seen in other Russian disinformation campaigns\*, often propagated by the use of bots. Also, since the comments from User B represent those held by the current authoritarian government it is counterintuitive that User A could be considered to be bullying. By its nature, bullying implies a power differential between someone more powerful harming someone who has less power. I do not understand how these online comments from a Russian dissident could be considered to inflict any real harm on an entity representing the views of, let's be blunt about it, Vladimir Putin. In my view, this is a case of perverting the Facebook community standard to silence political dissent.

Link to Attachment

No Attachment

2021-004-FB-UA

Case number

PC-10021

Public comment number

Europe

Region

Chris

Commenter's first name

Gray

Commenter's last name

English

Commenter's preferred language

DID NOT PROVIDE

Organization

No

Response on behalf of organization

----

Short summary provided by the commenter

The board is paying disproportionate attention to context, more than would have been available to any human making or reviewing the original decision. The comment in question was reported by the user it described, and made negative claims about their character. Deleting it is consistent with FB's policies on bullying.

Full Comment

I am a former content moderator for Facebook, and worked to make these kind of decisions at scale. Your description of the circumstances describes the original post, a response, and a response to that response. I don't know the current arrangements for moderators, but you may wish to ask Facebook what information would be available to anyone making a decision on this case. To the best of my knowledge, the original post made by user A would not be visible to anyone reviewing this case, they would only see User B's comment and A's response. And they would only see B's comment if A's reply was made using the reply function in response to it. ie a direct reply to B, not as a standalone comment in a thread even if intended as a reply. With this in mind, the amount of context you're considering is greater than would be available to a human moderator, possibly substantially so. If you want to evaluate the appropriateness of FB's response, it's important to do so using only the information that would be available to whoever made or reviewed the original decision. Regarding the decision itself, this is not a third-party report. User B has complained about the way they were described by User A, what we referred to as a "name or face match". The person being described is the one making the report, and their own feelings about the language used are taken into account. There is a higher threshold for acceptability, as a matter of policy. The use of the word "bot" may be problematic as it suggests the individual is not free-thinking. I would consider it similar to the original meaning of "drone" as a "mindless worker bee. It's a debatable term, one which FB is presumably clarifying for moderators. However the use of "cowardly" is a lot more clear cut. This is a negative character claim, just

like calling someone stupid, weak, evil, etc. If a user complains that someone has made a negative character claim about them, then that should require the removal of that comment. The context, or User A's intentions or justifications, would be irrelevant. Moderators are required to evaluate reported content on its own merits, without reference to the wider situation. I believe that FB's decision in this case is consistent with the policy as I remember it.

[Link to Attachment](#)

**No Attachment**

2021-004-FB-UA

Case number

PC-10024

Public comment number

Europe

Region

Tijana

Commenter's first name

Milosevic

Commenter's last name

English

Commenter's preferred language

National Anti-Bullying Research and Resource Centre

Organization

Yes

Response on behalf of organization

-----

Short summary provided by the commenter

Even though user B may have felt that they were bullied or harassed, we do not think that the case meets the criteria of bullying or harassment following the widely used definitions in research and practice with regard to bullying, cyberbullying and harassment; we do not think that the case meets Facebook's criteria for bullying and harassment either as specified in the Community Standards. Rather, this case appears to represent conflict between parties of equal capacity to defend themselves. There are no clear signs of repetition and power imbalance between the parties, which are typically the criteria necessary for a case to be considered as bullying/cyberbullying.

Full Comment

1. 1. Whether Facebook's decision complied with its Community Standard on Bullying and Harassment We would like to raise the issue of whether this case should have been treated as "Bullying and Harassment" in the first place. Bullying and harassment are two distinct concepts, and we do not think that this case necessarily meets the criteria of bullying and harassment. While there is no agreement on bullying definitions among scholars, and definitions may vary as the concept is applied in different settings (e.g. cyberbullying among children in a school context vs. among adults), the key criteria of bullying, and therefore cyberbullying, which is derived from the definition of offline bullying, are "intentionality" "repetition," and "power imbalance". Scholars tend to agree on these main features of bullying: 1) Intentionality: Bullying is a goal-oriented and systematic abusive behaviour (intention to hurt, an aggressive act); 2) Repetition: Bullying happens repeatedly and 3) Imbalance of power: The systematic abuse of power is perpetrated by someone who is either physically or psychologically stronger than the target. Bullying has a negative impact on the target, in terms of

their psychological and physical wellbeing. Offline bullying can take physical, verbal and relational forms (e.g., social exclusion, gossiping) and the latter two are manifested variously in online environments. Offline and online bullying are highly correlated; i.e., cyberbullying is an extension of traditional bullying as targets often face the same perpetrators offline and online (Görzig, A., & Macháčková, 2015; Smith, 2016). One broad and widely accepted definition of cyberbullying recognizes this phenomenon as “wilful and repeated harm inflicted through the use of computers, cell phones, and other electronic devices,” (Hinduja & Patchin, 2009, p. 5), and this definition would apply to social media as well. Likewise, there are many definitions of harassment, but it tends to involve “persistent and repeated course of conduct targeted at a specific person that is designed to and that causes the person severe emotional distress, and often the fear of physical harm” (Citron, 2015, p.2). Therefore, in both cyberbullying and harassment there is repetition and there is intention to harm, typically with some consequences. In case of bullying, there is also the additional criterion of power imbalance between the perpetrator and the victim. This power imbalance can manifest itself online in many forms –the perpetrator can leverage anonymity, have better digital skills to execute an attack or they can simply have more followers and their attacks get more visibility. In the case under consideration, in our view, based on the evidence provided, there is little indication that there is a power imbalance present. Both sides were directing insulting comments at each other without one distinctly being a victim: Both user A and user B directed offensive comments at each other. This is why, we do not think that these comments meet the criteria of bullying or harassment. Looking specifically at Facebook’s Community Standards, and the provided examples of bullying or harassment, you write: “Do not (...) target anyone maliciously by (...) “targeted cursing,” “negative physical descriptions” and “negative character and ability claims.” This would imply that insults you took down should be taken down, but in the absence of systemic repetition and power imbalance we would likely not consider this to be harassment or cyberbullying. Furthermore, your standards prohibit “Posting content about a violent tragedy, or victims of violent tragedies that include claims that a violent tragedy did not occur” (...) Otherwise paid or employed to mislead people about their role in the event.” If you consider Navalny to be a victim of poisoning (August 2020) and his supporters protesting as supporters of a victim, then this might imply, if considering the situation very broadly, that user B’s comment on user A’s protest activity might violate your standards too. Therefore, taken altogether, we do not think that the case should be interpreted under bullying or harassment policy. • Whether Facebook’s decision and Community Standard on Bullying and Harassment complied with its values and its human rights responsibilities, considering both the importance of free expression and the need to ensure scalable responses to harms caused by bullying and harassment. Having in mind the political situation in Russia and the role of Alexey Navalny, we do not think that the decision complied with your values and human rights responsibilities, especially in the context of scalable responses to harms caused by bullying and harassment. This type of interpretation of bullying and harassment in our view is not necessarily correct and it should not be applied to other cases.

References Citron, D. K. (2015). Addressing cyber harassment: An overview of hate

crimes in cyberspace. *Case W. Res. JL Tech. & Internet*, 6, 1. Görzig, A., & Macháčková, H. (2015). Cyberbullying from a socio-ecological perspective: a contemporary synthesis of findings from EU Kids Online. Hinduja, S., & Patchin, J. W. (2009). *Bullying Beyond the Schoolyard: Preventing and Responding to Cyberbullying*. Thousand Oaks, CA: Sage Publications (Corwin Press). Smith, P. K. (2016). Bullying: Definition, Types, Causes, Consequences, and Intervention. *Social and Personality Psychology Compass*, 10, 519-532

Link to Attachment

[PC-10024](#)

2021-004-FB-UA

Case number

PC-10025

Public comment number

Europe

Region

Andrei

Commenter's first name

Soldatov

Commenter's last name

English

Commenter's preferred language

the Center for European Policy Analysis (CEPA)

Organization

Yes

Response on behalf of organization

-----

Short summary provided by the commenter

The use of an expression "bot" is unacceptable, since it's used to exclude an interlocutor from a further debate. But the comment under consideration consists of an eye witness eyewitness report of a politically important event, with the real name backing it. Thus, the comment has an additional value of being an important contribution to the debate about the political events of significance in Russia's political history.

Full Comment

The case is extremely interesting. The nature of a political debate in Russian Facebook has several peculiarities one needs to remember. First, it's the one of very few places where the unrestricted debate could take place since all other places – traditional media and the like, -- are placed under some sort of government control, and Russian social media are heavily censored and surveilled. Second, for the reason mentioned above, those debates are conducted under constant barrage of attacks by armies of trolls, mostly pro-Kremlin, and anti-government in some cases. In Russia, a troll means a person who doesn't seek to change the opinion of the interlocutor or the audience but sows discord by posting inflammatory, extraneous, or off-topic posts with the deliberate intent of disrupting normal on-topic discussion by provoking readers into an emotional response. At the same time, the Russian society is strongly divided, with almost no proper or meaningful dialogue between the opposing parties. In many cases the debates on hotly disputed or politically sensitive issues, including the ones on Facebook, are conducted on two levels – a meaningful conversation among the people which mostly share the same set of values and views, and the constant search/detection/identification of the opposing force with whom any debate is meaningless with the goal to exclude him/her from a further conversation as soon as possible. The identifying and unmasking of trolls have been seen by many Facebook users in the country as a necessary precaution

measure, but it also became an easy way to deal with the opponent who challenged one's comments. The use of an expression "bot" – meaning essentially, dehumanization of an opponent as a robot on somebody payroll incapable of having his/her own views and opinions, -- became widely accepted as a means to exclude an opponent from a conversation. This, of course, cannot be welcomed. "Common cowardly bot" [банальный трусливый бот], the expression used by the User A, was meant, in my opinion, as a way to keep the User B out from a further conversation, which is against the very nature of what Facebook became for Russian users. The Facebook would have every right to take that comment down, if it were consisting of that expression only. However, the User A insists that in the comment under consideration the user "claimed to have participated in the Saint Petersburg protests with their colleagues and adult children. User A stated that there were so many protesters that they lost their companions in the crowd. They claimed to have witnessed elderly and disabled protesters." Taking that together with the real name used by Facebook users as a rule, that comment could be considered an eyewitness report of a politically important event, with the real name backing it. Thus, the comment has an additional value of being an important contribution to the debate about the political events of significance in Russia's political history. The comment which triggered the exchange between the User A and the User B took place next day after the protests held across the country in support of the opposition leader Alexei Navalny after his arrest. These protest rallies were met with brutality and excessive use of force by the anti-riot police units in many cities, including Saint Petersburg. It would be much welcomed if Facebook define a more differentiated response to users' complaints, and to use warnings messages more extensively. In that case, the warning from the Facebook administration to refrain from using the "bots" expressions in comments could be seen as a perfectly sufficient measure.

[Link to Attachment](#)

No Attachment



**2021-004-FB-UA**

Case number

**PC-10026**

Public comment number

**Europe**

Region

**GIOVANNI**

Commenter's first name

**PANSINI**

Commenter's last name

**English**

Commenter's preferred language

**Avv. Giovanni Pansini**

Organization

**No**

Response on behalf of organization

-----

Short summary provided by the commenter

**See documents (in red my comments)**

Full Comment

**See the doc (comments in red)**

Link to Attachment

[PC-10026](#)

2021-004-FB-UA

Case number

PC-10027

Public comment number

United States and Canada

Region

Clint

Commenter's first name

Watts

Commenter's last name

English

Commenter's preferred language

Foreign Policy Research Institute

Organization

No

Response on behalf of organization

-----

Short summary provided by the commenter

User A's content should not have been moderated and did not violate Facebook's policy regarding "Bullying and Harassment."

Facebook should reinstate User A's post based on three factors.

Full Comment

First, User A's use of the phrase "common cowardly bot" does not represent bullying. Having reviewed Facebook's Bullying and Harassment Community Standard, I do not see how the language of "common cowardly bot" can be assessed as a malicious violation threatening physical harm, depicting a violent tragedy or advancing sexual assault or abuse. The term "bot" is generic, suggesting inauthentic behavior indicative of computational propaganda, and there is no evidence a harm results from a social media persona being called a "bot." The use of the adjective "cowardly" is a term connoting an opinion—a term which is widely used in many cultures and contexts. While the entity deemed as cowardly may not receive the opinion favorably, the claim does not meet the threshold for "Bullying and Harassment" nor does the "cowardly" assertion arise repeatedly to suggest sustained harassment. Finally, there's no evidence to suggest that User A posted "content about victims or survivors of violent tragedies ... acting/pretending to be a victim of an event" or "otherwise paid or employed to mislead people about their role in the event." In contrast, User B's content claiming protestors were mentally "slow" and "used" more closely aligns with violative Bullying and Harassment content than User A's, but again, does not meet the threshold for removal from Facebook. Second, this case raises an additional contextual question: Did User A have reason to levy the claim that User B may be a "bot"? Based on the recent history of Russian state-backed inauthentic activity on social media, User A had reason to question the authenticity of the Facebook post. Dating back to as early as

2003, Russian security services and state-sponsored media entities have been identified as promoting totalitarian values and suppressing pro-democratic speech through organizations referred to by a host of names to include:

- Russia troll army (Армия троллей России),
- Russian bots (Русские боты)
- Putinbots
- Kremlinbots (Путиноботы, кремлеботы)
- Troll factory (Фабрика троллей)
- Lakhta trolls (Лактинские тролли)

The growth of social media platforms gave rise to another organization, known as the Internet Research Agency (IRA), which by 2013 had been documented as leading a smear campaign against Alexei Navalny—the same Alexei Navalny at the center of the protests to which User B’s comment refers. Internet Research Agency employees have also been documented posting comments in support of the Russian government while employing false personas. From 2017 through 2019, the U.S. Special Counsel Investigation into Russian interference in the 2016 U.S. election and the Senate Select Committee on Intelligence investigation into Russian Active Measures and use of social media documented how the Kremlin and its proxy organizations employ individuals, commonly referred to as “trolls,” and computational propaganda, often referred to as “bots,” to suppress its political opposition at home and abroad. Facebook has repeatedly shuttered Russian government-sponsored inauthentic activity emanating from the Internet Research Agency, and Twitter disclosed in January 2018 that they assessed more than 50,000 automated accounts (i.e. “bots”) linked to Russia were tweeting election-related content during the U.S. presidential campaign. User A had good reason to be suspicious of inauthentic “bot”-type posts against pro-democracy protests based on the pervasive history of the Russian government suppressing dissent via inauthentic social media activity. Third, beyond the Kremlin’s recent history of social media manipulation, User A likely questioned the authenticity of the post and its sentiment because User B’s narratives matched the Russian government’s current line at the time, which broadly sought to suppress and undermine pro-democracy protests occurring inside Russia in support of Navalny following his poisoning and subsequent detainment by Russian authorities. The narrative that Western social media companies were being used to manipulate children did not arise organically and while it cannot be proven from this single incident, coordinated messaging from social media users that matches the narratives of the Russian state has, in the past, been highly indicative of “troll” or “bot”-type activity. During the same time span in which this case occurred, a Meduza article from January 22, 2021 noted a flurry of anti-Navalny and anti-protest content published by various prominent Russian influencers leading up to the January 23, 2021 protest, planned in support of Navalny. Separately, allegations of coordinated inauthentic activity surfaced on January 25 when it was reported that Russian bloggers were allegedly offered payment for posting anti-Navalny content. Ridicule of the protest’s authenticity and claims of foreign manipulation inciting minors to protest were quickly pushed by Russian state-backed media and previously identified pro-Russian government propagandists like Maria Butina—the Russian national charged by the U.S. Justice

Department in July 2018 for “conspiracy to act as an agent of the Russian Federation Within the United States.” The content moderated in this case—while it does not meet the threshold for Bullying and Harassment—raises concerns that the flagging system for moderation of social media content may be used to delegitimize organic protests in Russia.

Link to Attachment

[PC-10027](#)

2021-004-FB-UA

Case number

PC-10033

Public comment number

United States and Canada

Region

Daniel

Commenter's first name

Gainor

Commenter's last name

English

Commenter's preferred language

Media Research Center

Organization

Yes

Response on behalf of organization

-----

Short summary provided by the commenter

**This is a silly attempt to adjudicate an ordinary disagreement. People must be allowed to disagree and even use strong language doing so.**

Full Comment

This is a ridiculously petty conflict to appeal to the board. Yes, ordinary people need to be allowed to disagree with one another about politics. And if the other user doesn't like it, they can block the person commenting. Problem solved. As longtime reporter Nat Hentoff famously stated, "[F]ree speech is to support the right to speech of people you hate." And indeed, the Founding Fathers of the United States would agree, the First Amendment to the United States Constitution qualifying that: "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble ... ." With the fundamentals of free speech laid out in the two quotations above, the Media Research Center proudly supports the principles of mirroring the First Amendment. The Facebook Oversight Board should afford its users nothing less than the free speech embodied in the First Amendment. The Board must not trample upon the free speech liberties that have long made the United States a beacon for freedom. Facebook suggests in its Community Standards that the platform "recognize[s] how important it is for Facebook to be a place where people feel empowered to communicate," but such communication cannot be had if free speech is not allowed on the platform. The Board should, as Facebook's Community Standards provide, "allow discussion, which often includes critical commentary."

Link to Attachment

No Attachment

2021-004-FB-UA

Case number

PC-10036

Public comment number

Europe

Region

Garry

Commenter's first name

Kasparov

Commenter's last name

English

Commenter's preferred language

Human Rights Foundation

Organization

No

Response on behalf of organization

-----

Short summary provided by the commenter

The enforcement was an error as User A's comment in no way qualifies as harassment or bullying. User B's post and subsequent complaint are typical of abusive behavior by Russian government agents against those supporting opposition activities in Russia.

Full Comment

15 March 2021 Dear Members of the Oversight Board: I submit this public comment from extensive personal experience with the issues identified by the Board in this case. It is very similar to cases involving my organizations and personal online properties. The very existence of this case is a good example of the methods of abuse employed by the Russian regime and its professional and amateur supporters. They are expert at manipulating the community standards and automated reporting tools at Facebook and other social media sites to attack the posts and accounts of those critical of the Russian government and its actions. In this case, those who post in support of the Russian regime's critics, such as Alexei Navalny and his supporters. • Facebook's decision to remove the first-hand report on the protests by User A is a clear error. The Russian phrase they used "банальный трусливый бот", is in no way a personal attack that should be called bullying or harassment. It's a mild descriptor, a term used for those who operate in an automated fashion (like a robot, or bot) in service of harassing critics of the Russian regime and spread misinformation. Which, based on the facts of the case, is exactly what User B was doing and how they were behaving. • The report by User B to have User A's factual response post removed is typical "bot" behavior. The post's removal in response to this frivolous complaint illustrates the failure of Facebook's enforcement mechanism to defend its users from coordinated state actor activities against dissent. • There is no freedom of speech in Russia today. All major media is under direct influence of the regime, and censorship and

persecution of any opposition activity is normal. My own news website is censored by the Russian authorities. Unable or unwilling to completely block social media sites like Facebook, the Russian method has been to use these “bot” techniques to spread propaganda and to abuse the enforcement mechanisms against any content or people they deem unfriendly to the regime of Vladimir Putin. • That User A’s post was removed, and that the Board has now spent so much time and effort to refute this absurd complaint is proof that the Kremlin’s campaign is successful. False and slanderous information was posted by the regime’s agent. A factual refutation by an opposition supporter was removed upon the complaint by the regime’s agent. I hope Facebook looks seriously at fortifying their enforcement mechanisms against these manipulations. The Russian opposition has enough to worry about without one of the few tools remaining to us to organize and express our dissent becoming yet another tool of the authoritarians. Regards, Garry Kasparov Founder and Chairman, Renew Democracy Initiative Chairman, Human Rights Foundation 13th World Chess Champion

Link to Attachment

[PC-10036](#)

2021-004-FB-UA

Case number

PC-10038

Public comment number

Europe

Region

Sultan

Commenter's first name

Suleimanov

Commenter's last name

English

Commenter's preferred language

Meduza

Organization

No

Response on behalf of organization

-----

Short summary provided by the commenter

The term "bot" is used by Russian Internet users not as an insult, but as an attempt to suppress meaningless discussion. In this context, "bot" is a synonym for the word "troll"

Full Comment

В российском интернет-пространстве существует класс «прокремлевских троллей» — пользователей, которые пытаются развязать бессмысленные дискуссии на политические темы, чтобы либо убедить других пользователей не поддерживать оппозицию и протесты, либо хотя бы лишить других пользователей желания участвовать в подобных дискуссиях и выражать собственное мнение. Таких пользователей иногда называют «ботами», хотя технически это не так — чаще всего это живые люди, которые занимаются разжиганием дискуссий за деньги, работая на организации, предположительно связанные с про-государственными структурами. В этом смысле, указание на то, что кто-то является ботом — это не попытка оскорбить пользователя, а скорее попытка закончить дискуссию, в которой невозможно переубедить оппонента (так как он не выражает собственное мнение). <https://www.nytimes.com/2018/02/18/world/europe/russia-troll-factory.html>

Link to Attachment

No Attachment



2021-004-FB-UA

Case number

PC-10040

Public comment number

Latin America and Caribbean

Region

João Victor

Commenter's first name

Archegas

Commenter's last name

English

Commenter's preferred language

Institute for Technology and Society of Rio

Organization

Yes

Response on behalf of organization

-----

Short summary provided by the commenter

Under the circumstances that were brought before this Board, it is our perception that the expression "common cowardly bot" should not be considered a violation of Facebook's bullying and harassment policy. The company's decision may unduly restrict an instance of legitimate speech and undermine its own values and human rights responsibilities. Moreover, the decision may not satisfy the "legality" test, which requires platforms to be more specific and transparent when restricting freedom of speech. By failing to point to a specific clause within the bullying and harassment policy and explaining its reasoning in this case, Facebook sets a precedent that limits speech without offering enough information for its users to accurately predict how

Full Comment

The matter at hand raises important preliminary concerns on the general understanding of what a bot is. According to the specialized literature on the subject, a bot - and more specifically a social bot - is a fully or partially automated account on social media, created with the purpose of making certain tasks automatic such as liking or sharing and, in fewer cases, commenting or executing other tasks on such platforms (Grimme, Assenmacher, Adam, 2018; Assenmacher et al, 2020). Many studies have confirmed the employment of this model in the dissemination of mis/disinformation campaigns, disrupting the harmony of public online debate, specially in political contexts including elections, protests and other situations of unrest (Bessi and Ferrara, 2016; Ferrara, 2017). The characterization of social bots and, consequently, their detection is still a debated issue among scholars and technologists, as their creation aims at imitating human online behavior. Therefore, the elements that allow one to state that a certain account or profile is a bot are not 100% accurate and cannot be taken for granted, as these characteristics

are constantly evolving and changing. In practice, the identification of an account or profile as a “bot” actually means that such an account displays elements of automated behavior and not that it is categorically a bot. With that in mind, it is our understanding that User A's comment referring to User B as a "common cowardly bot" should not be perceived as a violation of Facebook's bullying and harassment policy under the circumstances that were brought before this Board. The company's decision to remove the comment raises concerns that the platform may unduly restrict legitimate speech in the future, undermining its own values and human rights responsibilities. Moreover, Facebook's account that this specific piece of content was removed because it "degrades or shames a private individual" may not satisfy the "legality" test, which requires platforms to be more specific and transparent when restricting freedom of speech. In Case Decision 2020-003-FB-UA, regarding the use of the term "taziks" by an Armenian to describe Azerbaijanis during an armed conflict between the two nations, this Board acknowledged that "there may be instances in which words that are demeaning in one context might be more benign, or even empowering, in another". While in that case the conclusion was that during an armed conflict the use of the term "taziks" was "meant to dehumanize its target", we believe that in this case the circumstances call for a different conclusion. Although User A called User B a "common cowardly bot", the two adjectives on their own do not rise to the level of what may be considered bullying or harassment under Facebook's community standards. What really seems to be at issue in this case is the use of the term "bot" to refer to another person. Although the "taziks" case involved the enforcement of Facebook's hate speech policy, it is interesting to note that the bullying and harassment policy also prohibits the use of language meant to dehumanize another person through comparisons to animals, insects or inanimate objects. It is true that calling someone a "bot" amounts to, at first glance, dehumanizing speech. However, User A called User B a "bot" for questioning his first-hand experience during a protest in Russia in support of Alexei Navalny. It is our perception that, under similar circumstances to those of the case at hand, the comparison to an inanimate object is not meant to degrade or shame the interlocutor, but instead to legitimately suggest that one user is a piece of code employed to delegitimize another's personal account of a political event. As suggested above, the use of the term "bot" is a suggestion that a specific profile displays traces of automatic behavior, even if it is not emphatically a bot. In terms of human rights law, this Board rightly noted in Case Decision 2020-003-FB-UA that "Article 19, para. 2, of the ICCPR gives heightened protection to expression on political issues" and that this "protection remains even where those claims may be inaccurate or contested and even when they may cause offense". What is more, a decision to limit someone's right to freedom of expression on political issues, especially in places that are experiencing democratic backsliding, is inconsistent with Facebook's value of "voice". Although in some instances Facebook requires self-reporting by the person targeted by a specific content, such requirement is not present when "private or involuntary public figures" are targeted with "comparisons [...] to an inanimate object". As it stands, we fear that Facebook's decision may serve as a precedent to silence legitimate political speech when one user refers to another as a "bot", even if the claim is proven to be inaccurate. Finally, we believe that

Facebook's decision fails to satisfy the requirement of "legality" under Article 19, para. 3, of the ICCPR. Because it may amount to dehumanizing language, the word "bot" was very likely the chief factor behind the decision to remove User A's comment. However, the company's decision was not specific enough for us to be certain. As this Board noted in the "taziks" case, "individuals must have enough information to determine if and how their speech may be limited, so they can adjust their behavior accordingly". By failing to point to a specific clause within the bullying and harassment policy, Facebook sets a precedent that limits speech without offering enough information for its users to accurately predict how this rule will be enforced in the future. After all, as mentioned above, calling someone a "bot" may only mean that a specific account is displaying elements of automated behavior. It is important to stress that the "legality" requirement should not only apply to the way Facebook writes its rules, but also to the way it construes them.

Link to Attachment

[PC-10040](#)

2021-004-FB-UA

Case number

PC-10041

Public comment number

Europe

Region

Leonid

Commenter's first name

Volkov

Commenter's last name

English

Commenter's preferred language

Navalny's campaign

Organization

Yes

Response on behalf of organization

-----

Short summary provided by the commenter

I believe that the comment should be considered within the context of a discussion. Unfortunately, the debate Russian speaking Internet is often manipulated or even dominated by the government-sponsored bots and trolls, which operate using "темники", the "playbooks". Suggesting that someone is a bot is not really a personal offence, but rather a reflection of these sad reality.

Full Comment

During the protests in January 2021, which were related to the unlawful imprisonment of the leader of Russian opposition Alexey Navalny, just after he came back to Russia from Germany, where he was treated after a near-fatal poisoning with a chemical nerve agent Novichok, the government's narrative was very much concentrated on the alleged "misuse of minors for the sake of protest" by the protest movement. This allegation was never close to truth: never did the opposition movement call for minors to participate in the street rallies (moreover, we warned them against the possible consequences: <https://youtu.be/6HAGMSQQ7Fo>), and the number of minors among the participants was somewhere near the long-time average (around 5% according to various research groups). Still, the government kept pushing the fake agenda about "opposition leaders trying to hide behind schoolboys backs". Apparently, this was very insulting for any actual participant of the protest. First, because he or she saw, that this is just not true. Second, because he or she was indirectly accused of being a coward, trying to push schoolboys ahead of themselves. So the suggestion of the user B has actually been a severe insult towards user A, and user's A reaction was quite natural — as he saw that user B is pushing a fake agenda, very similar to that disseminated by the state propaganda and their bots, he supposed that the user B might have been a bot himself. Albeit apparently wrong, this assertion does not

look like reasonless. Altogether, I would say the discussion in question is very much within the widely accepted norms of political debate in Russia, given all the context.

[Link to Attachment](#)

**No Attachment**