



Public Comment Appendix for 2021-011-FB-UA

Case number

Case description

Note: To allow people to provide comments on the nature and impact of the post and help people understand the Board's eventual ruling in this case, we are sharing some of the exact words used in this post. We do so in the interest of transparency, while recognising that some of the quoted language has the potential to offend.

In May 2021, a Facebook user who appears to be in South Africa posted in English in a public group described as focused on unlocking minds. The post discusses "multi-racialism" in South Africa, and states that poverty, homelessness and landlessness have increased for black people in South Africa since 1994. It also states that white people hold and control the majority of wealth, and that wealthy black people may have ownership of some companies, but not control. The post then states that if "you think" sharing neighbourhoods, language and schools with white people makes you "deputy-white", then "you need to have your head examined". The post concludes with "[y]ou are" a "sophisticated slave", "a clever black", "'n goeie kaffir" or "House nigger".

The post received over 1,000 views and was shared over 40 times. The user's Facebook profile picture and banner photo depict black people (the Board is not able to verify the identity or protected characteristics of users who appeal or report content).

Facebook removed the post under its Hate Speech policy the same day it was posted, after it was reported by a user who appears to be located in South Africa. Under its Hate Speech Community Standard, Facebook takes down content that "describes or negatively targets people with slurs, where slurs are defined as words that are inherently offensive and used as insulting labels" on the basis of their race, ethnicity and/or national origin. Facebook also prohibits targeting people based on protected characteristics with generalisations about mental deficiencies or statements of inferiority. The Community Standard includes an exception to allow people to "share content that includes someone else's hate speech to condemn it or raise awareness" and to take into account that "speech that might otherwise violate our standards can be used self-referentially or in an empowering way".

The user submitted their appeal to the Board in English. The user stated in their appeal that they want to understand why the post was removed. They noted that people should be allowed to share different views on the platform and "engage in a

civil and healthy debate". The user also stated that they "did not write about any group to be targeted for hatred or for its members to be ill-treated in any way by members of a different group". They argued that their post instead "encouraged members of a certain group to do introspection and re-evaluate their priorities and attitudes". The user also stated that there is nothing in the post or "in its spirit or intent" that would promote hate speech, and that it is unfortunate that Facebook is unable to tell them what part of their post is hate speech.

The Board would appreciate public comments that address:

- Whether Facebook's decision to remove the post is consistent with the company's Hate Speech Community Standard, specifically the rules against describing or negatively targeting people with slurs and generalisations about mental deficiencies or statements of inferiority.
- Whether Facebook's decision to remove the post is consistent with the company's stated values and human rights responsibilities and commitments.
- The usage and impact of the words included in this post in the South African context, including in discussions relating to the political, economic and social issues raised by the user.
- Content moderation challenges specific to South Africa, both in terms of respecting freedom of expression and addressing harms that may result from hate speech online.
- How Facebook should interpret the following concepts when enforcing its Hate Speech Community Standard: 'self-referential', 'empowering', 'condemning' and 'awareness raising'.
- Whether sufficient detail is currently provided to people who use Facebook in English whose content is removed for violating the "Hate Speech" policy.
- What information about users, including on protected characteristics, should be available to moderators when reviewing content, considering its possible relevance to enforcing the Community Standard on Hate Speech. The Board would also appreciate comments on whether Facebook can confirm user-provided information, and any privacy concerns these points might raise.

In its decisions, the Board can issue policy recommendations to Facebook. While recommendations are not binding, Facebook must respond to them within 30 days. As such, the Board welcomes public comments proposing recommendations that are relevant to this case.



Public Comment Appendix for

2021-011-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.

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
2021-011-FB-UA

Case number

6

Number of Comments

Regional Breakdown

1 Asia Pacific & Oceania	0 Central & South Asia	0 Europe	0 Latin America & Caribbean
1 Middle East and North Africa	3 Sub-Saharan Africa	1 United States & Canada	

2021-011-FB-UA

Case number

PC-10161

Public comment number

Middle East and North Africa

Region

Yoav

Commenter's first name

Moran

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 desicion to remove this post is correct IMO.

Full Comment

I understand the the post was written in anger and frustration, but I do not see these as good enough reasons to allow exception for hate speech. This kind of violent language has been shown to move quickly in a social network and to increase the atmosphere of violence and hatred. If the poster would like to protest something, the platform allows to do that, no slurs are needed. As we have a saying in Hebrew: Don't piss to the well you're drinking from. Facebook is our well.

Link to Attachment

No Attachment

2021-011-FB-UA

Case number

PC-10187

Public comment number

Sub-Saharan Africa

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

While I think that the poster has a legitimate argument for the posting not being hate speech, they did use terms that are offensive both in context in South Africa, as well as, currently, globally. More over they could have made the same argument without using those terms. I therefore think that while this is not hate speech, it was appropriate for Facebook to remove the post. What is regrettable is that the poster was not given the option to repost the text without using either the k-word or the n-word.

Full Comment

While I think that the poster has a legitimate argument for the posting not being hate speech, they did use terms that are offensive both in context in South Africa, as well as, currently, globally. More over they could have made the same argument without using those terms. I therefore think that while this is not hate speech, it was appropriate for Facebook to remove the post. What is regrettable is that the poster was not given the option to repost the text without using either the k-word or the n-word. Another consideration is the racial identity of the poster which would, in context, affect how the speech was perceived and experienced. Similarly, the identity of the person who made the complaint. But Facebook cannot, and should not, in my view, take this into account as it would open several new cans of worms as they would have to establish and reveal the racial identity of both the poster, and the complainant. It seems to me here that the reason this is a difficult decision is that Facebook's community guidelines are not designed to be context specific and do not really allow for nuance. It is very hard to apply these guidelines with sensitivity to context, when they were not designed with such sensitivity in the first place. In summary, I think that while this is not hate speech, it was correct to remove the post but I think an option by which the poster would be asked to edit the

text to substitute the two offensive words (k-word and n-word) would have been a far better outcome than total removal.

[Link to Attachment](#)

No Attachment

2021-011-FB-UA

Case number

PC-10190

Public comment number

Asia Pacific and Oceania

Region

Anjalee

Commenter's first name

de Silva

Commenter's last name

English

Commenter's preferred language

**Melbourne Law School,
The University of Melbourne**

Organization

No

Response on behalf of organization

Short summary provided by the commenter

Facebook's decision to remove the post is inconsistent with the company's Hate Speech Community Standard ('Community Standard'). Regard should also be had to the harms of hate speech that Facebook seeks to respond to and mitigate. The harms of hate speech are functional rather than expressive harms, and content should be moderated with regard to what such speech does rather than merely what it says. The identity of speakers is relevant to this assessment where there is ambiguity. In consultation with experts, it is important for Facebook to review the Community Standard and the company's moderation processes in order to engage with some of these complexities and nuances around how hate speech and potential hate speech manifest.

Full Comment

Please see attached.

Link to Attachment

[PC-10190](#)

2021-011-FB-UA

Case number

PC-10191

Public comment number

Sub-Saharan Africa

Region

William

Commenter's first name

Bird

Commenter's last name

English

Commenter's preferred language

Media Monitoring Africa and Rea411

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

Context is critical in assessing hate speech. In the current format we don't have the content so cannot offer informed view. We have a working system that works within human rights principles and standards to assess a range of digital harms. We would Be happy to share more.

Full Comment

Hate speech is a critical issue for us in South Africa and it's limits are currently before our constitutional court. Our organization is one of the amicus in the court matter. In addition to that we have set up a structure called the Real411 (www.Real411.org.za) recognised by our government and independent bodies we have worked with the electoral commission in 2019 and will be working with them again this year to combat Digital harms in run up to elections. I mention these issues because it's important that to assess and make a more useful contribution on the specific case we would need to see the post and the date. The use of the k word is deeply problematic for us in south africa and so when used it requires a good detailed analysis. So our request would Be to see the full post. We would be happy not to share it but could then give informed input. We are happy to participate further should there be more information requested.

Link to Attachment

No Attachment

2021-011-FB-UA

Case number

PC-10192

Public comment number

Sub-Saharan Africa

Region

Judith/ Michelle

Commenter's first name

Geldenhuys/Kelly-Louw

Commenter's last name

English

Commenter's preferred language

University of South Africa

Organization

No

Response on behalf of organization

Short summary provided by the commenter

Context is less important when it comes to posts on Facebook. Racial slurs must be taken at face value and measured against the Facebook Hate Speech Standards and prevailing law, particularly the national law of the country. It is better to play on the safe side in case of uncertainty in order to protect the vast number of people who may be negatively affected by a post on social media. Certain words and phrases in the context of race are considered particularly heinous by South African courts. These words qualify as hate speech because of their historic and cultural associations. Posts containing these words in the context of race should always be removed, saved in cases that would fall within the exceptions in the Hate Speech Standards.

Full Comment

COMMENT ON FIRST FACEBOOK HATE SPEECH POST REMOVAL CASE IN SOUTH AFRICA Background The post under scrutiny was posted on a public group on Facebook. The focus area of the group is unlocking minds and to discuss 'multi-racialism' in South Africa. The post was posted on Facebook during May 2021. The post was viewed 1000 times and was shared 40 times despite the fact that it was only on Facebook for a short period. The post was reported to Facebook by another user, i.e. a member of the group for contravening the Facebook 'Hate speech Community Standards'. Facebook removed the post on the same day. The user that posted the comment (the poster) appealed the decision to remove the post claiming that it did not qualify as hate speech. The Facebook Hate Speech Community Standards prescribes that content that 'describes or negatively targets people with slurs that are 'inherently offensive and used as insulting labels on the basis of their race, ethnicity and/or national origin'. Hate speech in context In order to protect people against hate speech and racist slurs different fundamental rights must be weighed

up: the rights to equality and dignity must be weighed against the right to freedom of speech. Factors considered in balancing these roles are: who the victim is, who the perpetrator is and the nature of the expression. Additional factors include: 'historical associations; who the utterer is as against the victim(s); the audience that is addressed and where the utterance is made; and the social conditions at the time of making of the utterance' (Geldenhuys J., & Kelly-Louw M. (2020). Demystifying Hate Speech under the PEPUDA. Potchefstroom Electronic Law Journal, 23, 1-50. <https://doi.org/10.17159/1727-3781/2020/v23i0a7520>). However, it is inconceivable that the reader of a social media post on Facebook or the Oversight Board, if the post is reported by someone, will have access to all of these background facts and be able to assess whether, in the legal sense, a posting qualifies as 'hate speech'. The person who reports a post would usually do so if he or she is offended by it, finds it inappropriate in the social and ethical context, and if in his or her mind the post is a serious contravention of what is understood under the relevant provisions in the Facebook Hate Speech Community Standards and/or the law. It is, therefore, our opinion that in assessing the desirability of the post on Facebook, and whether it ought to be removed, that the utterance must be taken 'at face value' and measured against the Facebook Hate Speech Community Standards provisions and the prevailing international and national law regulating hate speech and the making of derogatory statements or racial slurs. Given the number of potential victims that may be affected by a posting that does qualify as hate speech by any definition, it is our opinion that, in circumstances where uncertainty arises as to whether a post should remain or be removed based on alleged hate speech, it is safer to err on the safe side. The content of the post The statements that 'Poverty, homelessness, and landlessness have increased for black people in South Africa since 1994', that 'White people hold and control the majority of wealth' and that 'Wealthy black people may have ownership of some companies, but no control' hold at least some truth. These statements are neither offensive nor derogatory. What follows upon these statements is problematic: 'If you think sharing neighborhoods, language, and schools with white people makes you deputy white you need to have your head examined. You are a 'sophisticated slave', 'a clever black', "n goeie kaffir" or 'house nigger'. This statement contains words that are internationally and nationally unacceptable to utter in the context of race as was done here. The words that are used all bear cultural or historical associations that qualify them as hate speech. The word 'kaffir' as used in the post under scrutiny is always considered by South Africans to be hate speech, even if the perpetrator (the poster in this case) is black. (See Geldenhuys J., & Kelly-Louw M. (2020). Demystifying Hate Speech under the PEPUDA. Potchefstroom Electronic Law Journal, 23, 1-50. <https://doi.org/10.17159/1727-3781/2020/v23i0a7520> 23 and the national and international cases cited there). In our view the last part of the posting is a contravention of the Facebook Hate Speech Community Standards as it 'describes or negatively targets people with slurs that are 'inherently offensive and [is] used as insulting labels on the basis of their race, ethnicity and/or national origin'. Factors mitigating against a finding of hate speech and removal of the post The exceptions in the Facebook Hate Speech Community standard are that people may 'share content that include someone else's hate speech to condemn it or raise awareness'

and cite even hate speech in a 'self referential or empowering way'. These exceptions do not apply in this instance. It is not clear from the post that the poster was citing someone else's statements. What is apparent is that the poster condemns white people for the remnant inequalities in the South African community. He or she also denigrates the thinking, the mentality and standing of middle and upper middle-class black people who speak the same language as white people, whose children attend schools with white learners and who work together with and live alongside white South Africans. The poster contends that the intention was not to propagate hatred, but to 'encourage members of a certain group to do introspection and re-evaluate their priorities and attitudes' and not to 'target hatred for its members to be ill-treated in any way by members of a different group'. Instead, the poster claims that the intention was 'to share different views and engage in a civil and healthy debate'. However, the intention to stimulate a debate is not clear from the post at face value. If the idea was to enter in

Link to Attachment

[PG-10192](#)