

 Oversight Board

Public Comment Appendix for

Mention of the Taliban

in news reporting

Case number

Case description

In January 2022, the Facebook page of a news outlet in India shared a post with a link to an article on the news outlet's website. The post, in Urdu, contains text and an external link that leads to an article on the news outlet's website. The text states that Zabiullah Mujahid (the Taliban government in Afghanistan's Culture and Information Minister and official central spokesman) said that the Afghan New Year begins on March 21 and that schools and colleges for girls and women will open this year from the beginning of the new year. The article discusses this announcement in further detail. The page has about 14,000 followers.

A user initiated a report on the content to Meta, but did not complete their report. This incomplete report triggered a classifier that scored the content as potentially violating under the [Dangerous Individuals and Organizations](#) ("DIO") policy, sending the content for human review.

Meta removed this content for violating its Dangerous Individuals and Organizations Community Standard, having determined that it violated its prohibition on praising a designated terrorist group. Zabiullah Mujahid is a prominent member and spokesman for the Taliban, and the Taliban is a Tier 1 designated terrorist organization under Meta's DIO policy.

According to the Facebook Community Standards, praise of a designated entity includes “[s]peak[ing] positively about a designated entity or event;” “[giv[ing]] a designated entity or event a sense of achievement;” “[l]egitimizing the cause of a designated entity by making claims that their hateful, violent, or criminal conduct is legally, morally, or otherwise justified or acceptable;” and “[a]lligning oneself ideologically with a designated entity or event.” Meta states that it allows content which references dangerous individuals and organizations in the context of reporting on them, but “users must clearly indicate their intent when creating or sharing such content” and “if a user’s intention is ambiguous or unclear,” it will default to removing content.

The user who created the content appealed the removal, but Meta upheld its decision to remove the content. The user then appealed to the Oversight Board. When the Oversight Board brought the content to Meta's attention, Meta determined that this was an enforcement error and that this content fell into the

DIO policy exception for reporting and should not have been removed. Meta stated that it did not have any information on why this content was assessed twice as praise and not as news reporting.

In their statement to the Board, the user states that they are a media organization and do not support extremism. They say that their articles are based on national and international media sources and that this content was shared to provide information about women's and girls' education in Afghanistan.

The Board would appreciate public comments that address:

- How Meta's content moderation policies and practices affect public discourse about the Taliban's role in Afghanistan.
- How Meta's content policies and practices on Dangerous Individuals and Organizations affect the ability of journalists to report on these groups.
- The DIO policy prohibition on "praise" of Tier 1 and 2 designated individuals and entities and its compatibility with Meta's human rights responsibilities.
- What principles should guide if or when Meta should revoke or change the designation of an entity under the Dangerous Individuals and Organizations Community Standard, including for entities that form or take the place of governments.
- Whether Facebook's Dangerous Individuals and Organizations Community Standard unnecessarily limits discussion of designated groups that either form or take the place of governments, including in relation to "false positive" removals of media reporting and other commentary on current affairs.
- The relationship between US law prohibiting material support of designated terrorist organizations and Meta's content policies, and how this may affect freedom of expression globally.

In its decisions, the Board can issue policy recommendations to Meta. While recommendations are not binding, Meta must respond to them within 60 days. The Board welcomes public comments proposing recommendations that are relevant to this case.

 Oversight Board

Public Comment Appendix for

Mention of the Taliban

in news reporting

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.



Oversight Board

Public Comment Appendix for

Mention of the Taliban

in news reporting

Case number

6

Number of Comments

Regional Breakdown

1	0	4	0
Asia Pacific & Oceania	Central & South Asia	Europe	Latin America & Caribbean
0	0	1	
Middle East and North Africa	Sub-Saharan Africa	United States & Canada	

Mention of the Taliban
in news reporting

PC-10436

Case number

Public comment number

United States and Canada

Kaylana

Commenter's first name

Mueller-Hsia

Commenter's last name

Region

Brennan Center for Justice

Organization

English

Commenter's preferred language

Yes

Response on behalf of organization

Short summary provided by the commenter

Facebook's DIO standard has proven unworkable. Its reliance on vague terms like "praise" and "support" suppress critical political discussion and disproportionately affect minority communities and the Global South. And Meta's automated tools time and again fail to account for context, particularly in languages other than English. The board should recommend that Meta shift its focus from so-called dangerous groups and individuals to violent content and ensure that it devotes sufficient resources to content moderation in all the countries and languages that it operates.

Full Comment

See attachment--requirement on case page says no more than 5 pages, but this form still limits the submission to the original shorter requirement of 2 pages.

Link to Attachment

[PC-10436](#)

Mention of the Taliban in news reporting

PC-10435

Europe

Case number

Public comment number

Region

Charlotte

Baxmann

English

Commenter's first name

Commenter's last name

Commenter's preferred language

DID NOT PROVIDE

Yes

Organization

Response on behalf of organization

— — — —

Short summary provided by the commenter

This comment addresses questions concerning the DIO policy prohibition on “praise” of Tier 1 and 2 designated individuals and entities and its compatibility with Meta’s human rights responsibilities as well as the question which principles should guide if or when Meta should revoke or change the designation of an entity under the Dangerous Individuals and Organizations Community Standard, including for entities that form or take the place of governments. It was written by a group of law students of the Humboldt-University of Berlin on the occasion of a project on Digitalization and Institutions, especially addressing the issue of content regulation. We believe that the content should be restored.

Full Comment

On the question does Facebook's Dangerous Individuals and Organizations Community Standard unnecessarily limits discussion of designated groups that either form or take the place of governments, including in relation to "false positive" removals of media reporting and other commentary on current affairs? Facebook's Dangerous Individuals and Organizations Community Standard means that there is always the possibility that discussions about certain groups that either form governments or take their place will be restricted. This is due to the fact that the proper subsumption of a post under the DIO can be difficult or impossible in borderline cases, and as a result posts that appear to be in violation of the standard, but are actually made with other intentions, have to be deleted if doubts are left. To deal with this problem, the DIO provides for users to post content on the platform that could allegedly violate the DIO just if they must make their intention clear. It also reserves the right to META to delete the amount in case of doubt if the intention is not clear. However, the mere restriction of the very broad deletion order by the user clearly stating his intention is not sufficient to exclude cases in which information is passed on without being part of a direct discussion and the user is

not likely to even think about writing his intention. In general, it is to be supported that a restrictive approach is taken here, which is why the rule of doubt is correct, which orders deletion if the purpose of the post is unclear. Only a restrictive approach to more information-heavy posts can ensure that information provided by terrorist or similar organisations does not take on the status of that provided by governments. This danger would exist, however, if even in discussions reference were made to information from such organizations, since a discussion that is held online, unlike a discussion in the presence of all persons, is directly written down, and the arguments made here could build the basis of information for third parties, who could draw false conclusions from it. At the same time, however, there are certain topics of general information value where the description of one's own intention is unrealistic. The present case shows well that there is no promotion of the organization's information here, let alone a sharing of its ideology. It would also be absurd to question the position of the Taliban when announcing when women will actually have the opportunity to go to school again due to the current conditions in Afghanistan. The concrete situation in Afghanistan shows that the DIO does not take sufficient account of the fact that information and statements from a terrorist or similar organization also become informational as soon as they actually exert power. Then the ideologised passing on of infused statements becomes pure capitulation to the actual circumstances. In this respect, the DIO must be extended to include the criterion of social relevance. It is therefore necessary to dispense with a precise statement of intent if the information passed on by the user contains information that continues to be of relevance to certain sections of society. This is first of all the case if, due to certain power relations, the statements of the organizations have a decisive influence on the actual events. Furthermore, the requirements for the information content must be precisely defined. For example, fundamental rights could be used as a guideline, and circumstances that affect the content of fundamental rights could be defined as "socially relevant". This would also include access to education, which can actually be determined by the Taliban, so that it is both a fact and a matter of fundamental rights. Thus, the DIO is currently too restrictive, unnecessarily limiting the discussion on such groups. At the same time, adding the criterion of reflecting real power relations and information of social relevance in the area of fundamental rights would remedy this.

Link to Attachment

No Attachment

Mention of the Taliban in news reporting	PC-10432	Europe
Case number	Public comment number	Region
Withheld	Withheld	English
Commenter's first name	Commenter's last name	Commenter's preferred language
Withheld		Yes
Organization		Response on behalf of organization

Short summary provided by the commenter

This comment addresses the question which principles should guide if or when Meta should revoke or change the designation of an entity under the Dangerous Individuals and Organizations Community Standard, including for entities that form or take the place of governments. It was written by a group of law students of the Humboldt-University of Berlin on the occasion of a project on Digitalization and Institutions, especially addressing the issue of content regulation. We believe that the content should be restored.

Full Comment

The term "non-state actor" chosen by Meta in the Dangerous Individuals and Organizations policy under "TIER 2" can lead to issues, especially when a terrorist group temporarily or even completely takes over leadership of a state and acts as a "government". An interpretation of this term must thus be found that both makes possible extensive media coverage of such a group and still prevents the spread of pro-terrorist propaganda. The international recognition of a terrorist organization as a regular government as well as that organization's institutional consolidation in the country in question could serve as effective criteria to guide Meta's classification. Even if recognition of a foreign government is not constitutive for its qualification as "government", this criterion can be helpful when it comes to public interest and the form that media coverage should take when it comes to reporting the actions of a terrorist entity posing as a government. Nevertheless, international recognition of a government could be an arbitrary instrument of distinction. Even without recognition, a government can effectively rule over a country. This spikes a certain interest in media coverage, and hence requires the possibility to report on certain actions in a way that could be construed as positive. This is currently forbidden under the Dangerous Individuals and Organizations policy. A complementary criterion to balance out this issue could be the institutional

consolidation of the concerned organization in the country. As for the case of the Taliban, many countries linked their official recognition of the Taliban-government to certain conditions, like the respect of human (especially women's) rights, without admitting the effective control the Taliban are exercising over Afghanistan. Hence, there is a certain interest to allow a larger range of media reporting on the Taliban's actions, especially when it comes to actions that take steps towards the conditions posed by other countries since this might influence the official recognition of the Taliban by other states.

Link to Attachment

No Attachment

Mention of the Taliban
in news reporting PC-10431 Europe

Case number Public comment number

Region

Withheld **Withheld** **English**

Commenter's first name

Commenter's last name

Commenter's preferred language

Withheld

Organization

No

Response on behalf of organization

Short summary provided by the commenter

This comment addresses questions concerning the DIO policy prohibition on “praise” of Tier 1 and 2 designated individuals and entities and its compatibility with Meta’s human rights responsibilities. It was written by a group of law students of the Humboldt University of Berlin.

Full Comment

3. The DIO policy prohibition on “praise” of Tier 1 and 2 designated individuals and entities and its compatibility with Meta’s human rights responsibilities According to its DIO criteria, Meta deletes posts in which dangerous individuals or organizations are being “praised”, differentiating between Tier 1 and Tier 2 entities: While only content involving the “praise” of violence Tier 2 entities engage in is deleted, a post on Tier 1 entities can be deleted whenever it contains “praise” of the entity or a designated event. The question in this context must be whether the deletion of content that Meta qualifies as “praise” of a Tier 1 or Tier 2 entity violates the author's Freedom of Speech (which also protects the sharing of information through media) stipulated in Article 19 of the Universal Declaration of Human Rights. In its Corporate Human Rights Policy, Meta commits itself to the respect of human rights as stipulated in the International Bill of Human Rights, taking into account its responsibility as an important platform for the exchange of information. The deletion of a post infringes the author's Freedom of Speech. However, Article 19 (3) of the International Covenant on Civil and Political Rights allows for restrictions necessary to respect the rights of others or for the protection of national security, public order, public health or morals. The deletion could therefore be justified with Meta’s DIO policy serving the purpose of protecting the public order and morals by keeping Tier 1 and 2 entities from using Meta as a platform to promote themselves and their actions. It should be positively highlighted that Meta describes in detail which groups fall under the Tier 1-3 categories and which actions constitute “praise”, including examples. This could serve as an indication for the justification

of potential Freedom of Speech infringements by the reasoning of the post containing “praise” according to DIO policy and thus going against public order and morals. However, the distinction between “praise” and the objective reporting on “positive” events remains fluid. Since with Tier 2 entities, the “praise” needs to concern violent actions and the likelihood of reports to use positively connotated phrasing in the context of violence is rather low, this issue mainly concerns reports on Tier 1 entities. For posts linked to Tier 1 entities (which can concern further topics, A/M), however, there is an undeniable risk that objective reports might be classified as unpermitted content due to the low requirements set by Meta for content to be qualified as “praise”: The definition and examples provided leave one under the impression that posts (1) mentioning Tier 1 entities while (2) using positively connotated words like “brave”, “nothing wrong”, “I stand with” could be deleted regardless of whether in their text, the author shows support for the Tier 1 entity. The combination of aspects (1) and (2) could act as a trigger for deletion without further contextualization. This becomes clear in the present case: The reporting on the positively connotated event of girls being allowed to go to school combined with the mentioning of the Taliban’s spokesman triggered the deletion of the post even though the article did not contain signs of support towards the Taliban as such. While Meta states that it recognizes that users might share content on Tier 1 or 2 entities for informational purposes, in its policy it places an additional burden on the users by requiring them to clearly indicate their intent to avoid deletion. With this, Facebook transfers its own responsibility to prevent violations of the Freedom of speech regarding reports on DIOs, especially Tier 1 entities, to users. Instead of addressing the challenges presented by the complexity of the DIO policy prohibitions and the diversity of posts, instead of increasing the requirements for content to be qualified as “praise” or finding other ways to ensure the free objective reporting on DIOs, Facebook accepts the potential violation of its users’ Freedom of speech and makes it difficult for the users to make use of their rights. Hence, its DIO policy prohibition on praise, at least for Tier 1 entities, is incompatible with the human rights responsibilities it has committed itself to.

Link to Attachment

No Attachment

Mention of the Taliban
in news reporting

PC-10429

Europe

Case number

Public comment number

Region

Withheld

Commenter's first name

Withheld

Commenter's last name

English

Commenter's preferred language

Withheld

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

The commentary addresses the influence of Meta's content policies in news reporting and journalism, as well as public discourse about news from Afghanistan/about the Taliban.

Full Comment

The first question addresses whether Meta's content moderation policies affect the public discourse about the Taliban's role in Afghanistan - a complicated, political matter. Since its inception in 1994, the life for all the people living in Afghanistan - men and women, but especially women - has drastically changed. Women in Afghanistan have been given lesser access to education and the enjoyment of their rights. In the early 2000s, there was a small period of time with changes which gave us hope but since the Taliban offensive in 2021 and the following takeover, not many fundamental rights and securities were established and granted. Especially girls and women feared more than just interference with their fundamental rights and were left with nothing but fear and instability. There is no free press left in Afghanistan and everything is controlled and surveilled by the Taliban. Meta, a platform with a big global community, has the community goal to create a place for expression in which everyone is given a voice. Safety and dignity are few of the community standards. No content, which may harm the physical security of persons or disrespect the dignity and rights of others should be published and contributed on the platform and social media in general. Because of the past events, the Taliban are classified as a Tier 1 designated terrorist organization under Meta's DIO policy and Taliban-related content is often directly reported and removed. But what does it mean for the complicated matter of news reporting around the world? On the one hand, it is completely understandable to restrict all Taliban-related content, especially since their content is likely related to violence or content with need of a trigger-warning. On the other hand, it is important to consider the distinction between content "from" Taliban and content about the

Taliban. In this case here, the Indian news outlet created and released some content about the Taliban with the aim of informing consumers and spreading the news. Such content is important for the opening of a public discourse or critical questioning. To come to a conclusion: the content about the Taliban - whether positive or negative - is really difficult to disseminate on Metas platforms because the Taliban are directly classified as a Tier 1 designated terrorist organization under Meta's DIO policy. However this distinction is indispensable, since there is always a need for (global and international) news reporting and not every Taliban-related content is automatically „bad“ (especially since the world is not just designed in black and white and there is no morally „right and wrong“ or „good and „bad“). So yes, sometimes meta's content policies and practices affect public discourse about the Taliban's role in Afghanistan, since meta does not always differ in the types of Taliban-related content. This leads to missing information, missing critical questioning (etc...) which can influence the public discourse. The second question addresses whether meta's content policies affect the journalists to report on these groups. About the concrete possibly affected policies of Meta, see above. Meta's practices on Dangerous Individuals and Organizations are regulated in their DIO policy. Journalism and freedom of speech do not work without a stable news service and the lack of fake news. But what about it in this case? Usually, you should consider the granting of more women's rights as positive, but here, instead, the content was reported and removed. Of course, mentioning the Taliban, which is classified as a Tier 1 designated terrorist organization under Meta's DIO policy, is a possible violation and could warrant the suspension. But on the other hand, the Indian news outlet took a neutral position and its only aim was to spread the information, maybe even a small „message of hope“. In this case, it is as a matter of course something „good“, that girls and women can return to schools and continue their education. Although that possibility is limited to students up until the sixth grade, it still is nevertheless better than no education at all. In my opinion, it is still important to inform about such small, but nevertheless „positive“ changes. This kind of information can not only motivate those affected to act, it can also stimulate political discourse and encourage critical thinking all around the world and in meta's global, international community. Meta's content moderation policies and practices are surely a useful way to protect against dangers on the (hard to control) Internet, but this content moderation should not stand in conflict or even restrict the freedom of the press and information. Especially because the news outlet's Facebook page here did not take a biased position, the article was just informing about the developments in Afghanistan. An all-encompassing, neutral news coverage is not possible if content cannot be released or distributed at all. It is necessary to distinguish between the origin and destination of the content thoroughly and according to the individual case. Coming to a conclusion, in my opinion Meta's content policies and practices are affecting the ability of journalists, since the content in this case was e.g. removed, although it did not necessarily violate meta's content policies.

Link to Attachment

No Attachment

Mention of the Taliban
in news reporting

PC-10426

Asia Pacific and Oceania

Case number

Public comment number

Region

Gautam

Mishra

English

Commenter's first name

Commenter's last name

Commenter's preferred language

The University of Sydney

No

Organization

Response on behalf of organization

Short summary provided by the commenter

I believe that Meta's decision to remove the mention of the Taliban in reporting was wrong. Meta, although its responsibility is not to attempt to correct the ways of governmental organisations, is inherently apart of the information dissemination stream, serving as a the middle-ground to news organisations informing individuals. Should positive information surrounding otherwise negative actors come out, and there is no appraisal of otherwise inappropriate actions, individuals have a right to know what these actors are doing. Maintaining a historical profile of news organisations may make computational decisions of action against accounts would be beneficial. Articles could be checked against old articles for early intervention.

Full Comment

I believe that Meta's decision to remove the mention of the Taliban in reporting was wrong, and that the article and content like it should be kept on Meta's platforms. The Taliban allowing educational opportunities for girls and women is - to the vast majority of people - a good thing. It is also clear that the Taliban government has completed horrendous crimes against innocent civilians that have terrorised a nation and the world more over - and although slightly more contentious than the last statement - it too is accepted by the vast majority of people. 1: Fundamentally, the case is evaluating whether we should permit the publicising of a bad actor doing good things. In smaller cases than the Taliban, positive reinforcement can alter an actor's actions. I think that we should allow otherwise terms & conditions-breaking actors potentially more traction when they are not actively breaking the rules that bind them (say hate-speech or privacy related issues). This would be in an effort to create long-term notable changes in how otherwise problematic individuals' use social media, and could be implemented in smaller scale situations. Obviously this is not going to be the case with the example of the Taliban - no governmental

organisation cares for the approval of users on a social media network - but I don't believe correcting the ways of a governmental organisation is the role of Meta. However, Meta is inherently apart of this process. It is the role of journalists, opinion writers, and those involved in the media industry to criticise the institutions around us, be their criticisms approving or disapproving. This is then to be debated by individuals and, should they see something as deserving of criticism or change, individuals are the ones that will ultimately bring these criticisms and changes to fruition. Meta, being a platform for the media industry and individuals to interact, plays a key part in that information dissemination "stream" - being the way that these news organisations get to their individuals. It is widely known that the Taliban is a harmful institution, and news talking about the Taliban in a negative light does exist and should be viewable by users on Meta's platform. Under these specific circumstances, I believe that the argument that this will suddenly change individual's perceptions of the Taliban does not hold. I understand that this will not be the case of less talked-about harmful institutions, where there may be less negative coverage to allow for any amount of positive coverage. This is outside of the scope of this case - for this case and cases similar to it in size, this positive discourse, should it be correct, should be permitted onto Meta's platforms.

2: The glaring issue with this line of logic is where we should draw the line. In this case, I think there is a clear distinction between a bad actor doing good things. As such, it is easy to say that the article should be published - this is something that individuals should read to hold opinions on and enforce change upon the widely known to be harmful institution that is the Taliban. Should the article have other appraisals of inhumane actions the Taliban has committed, the onus back onto Meta. Meta ideally, although I understand this to be computationally intensive, would analyse not just the contents of the article linked, but maintain historical profiles of each of the news sources that have articles both reported and not reported on Meta's platforms. History is the best indicator of future actions. Should one of these news sources have a history of going against terms & conditions of Meta's guidelines, they should be limited in their ability to further promote material regarding the content that they have been flagged against. History will be the best indicator (alongside the face value content of the article) as to whether a publication should be regulated and where we should draw the line.

Should the article contain both promotion and disapproval, or have other articles posted later regarding other acts that are negative in the eyes of Meta, Meta could computationally check the news source against its historical profile. Should there be a pattern in their posts approving of these actions, then Meta could computationally regulate the post on its platforms. This could be completed in the form of a shadow ban or temporary ban of access and publicity to the account. These historical profiles should be public, and should be available for users to see when they anticipate clicking on a link to an article that would lead them down a path against Meta's regulations. It should be made clear to the individuals and organisations too that have received a historical profiling of their previous posts on the platform and off the platform, and that they may have access or features relating to their account be restricted in some way. This will increase the likelihood of an individual or an organisation to change their ways and continue using a Meta

platform. Again, as Meta systems exist already, the news source should be able to challenge these post and account restrictions.

There are obvious secondary issues that will be raised by this solution - issues relating to privacy, the result of mis-profiling, lost potential revenue to media institutions, and so on. However I believe that historical profiling and an examination of the core content of the article are the two best ways to computationally regulate similar situations to this from occurring again. Human verification will be needed from time to time, and I believe that to remedy issues at least surrounding mis-profiling and undue losses of revenue due to mis-profiling or post and account restrictions. Although monetarily and individually taxing, I don't think there's a future of moderation where the ultimate need for human intervention should it be necessary.

[Link to Attachment](#)

No Attachment