



Public Comment Appendix for

2022-007-IG-MR

Case number

Case description

In January 2022 an Instagram account which describes itself as publicising British music posted a video with a short caption on its public account. The video is a 21 second clip of the music video for a UK drill music track called "Secrets Not Safe" by the rapper Chinx (OS). The caption tags Chinx (OS) as well as an affiliated artist and highlights that the track had just been released. The video clip shows part of the second verse of the song and fades to a black screen with the text "OUT NOW." Drill is a subgenre of rap music popular in the UK, with a large number of drill artists active in London.

Shortly after the video was posted, Meta received a request from UK law enforcement to remove content that included this track. Meta says it was informed by law enforcement that elements of it could contribute to a risk of offline harm. The company was also aware that the track referenced a past shooting in a way that raised concerns that it may provoke further violence. As a result, the post was escalated for internal review by experts at Meta.

Meta's experts determined the content violated the Violence and Incitement policy, specifically the prohibition on "coded statements where the method of violence or harm is not clearly articulated, but the threat is veiled or implicit." The Community Standards list signs that content may include veiled or implicit threats. These include content that is "shared in a retaliatory context," and content with "references to historical or fictional incidents of violence." Further information and/or context is always required to identify and remove a number of different categories listed at the end of the Violence and Incitement policy including veiled threats. Meta has explained to the Board that enforcement under these categories is not subject to at-scale review (the standard review process conducted by outsourced moderators) and can only be enforced by Meta's internal teams. Meta has further explained that the Facebook Community Standards apply to Instagram.

When Meta took the content down, two days after it was posted, it also removed copies of the video posted by other accounts. Based on the information they received from UK law enforcement, Meta's Public Policy team

believed that the track “might increase the risk of potential retaliatory gang violence,” and “acted as a threatening call to action that could contribute to a risk of imminent violence or physical harm, including retaliatory gang violence.”

Hours after the content was removed, the account owner appealed. A human reviewer assessed the content to be non-violating and restored it to Instagram. Eight days later, following a second request from UK law enforcement, Meta removed the content again and took down other instances of the video found on its platforms. The account in this case has fewer than 1000 followers, the majority of whom live in the UK. The user received notifications from Meta both times their content was removed but was not informed that the removals were initiated following a request from UK law enforcement.

In referring this matter to the Board, Meta states that this case is particularly difficult as it involves balancing the competing interests of artistic expression and public safety. Meta explains that, while the company places a high value on artistic expression, it is difficult to determine when that expression becomes a credible threat. Meta asks the Board to assess whether, in this case and more generally, the safety risks associated with the potential instigation of gang violence outweigh the value of artistic expression in drill music.

The Board would appreciate public comments that address:

- The artistic and cultural significance of UK drill music and any relationship between the sharing of this music online and acts of real-world violence.
- Meta’s human rights responsibility to respect artistic expression, as well as to ensure its platforms are not used to incite violence, and how this should inform its approach to content moderation of music.
- Whether Meta’s policies on violence and incitement should include allowances for humour, satire, or artistic expression, and if so, how they should be worded, and how they should be enforced accurately at scale.
- How social media platforms should manage law enforcement requests for the review or removal of content that does not violate national laws but may breach a platform’s content rules.
- How social media platforms should incorporate law enforcement requests for content removal, especially requests not based on alleged illegality, into their transparency reporting.

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



Public Comment Appendix for
2022-007-IG-MR

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

2022-007-IG-MR

Case number

10

Number of Comments

Regional Breakdown

0	2	1	1
Asia Pacific & Oceania	Central & South Asia	Europe	Latin America & Caribbean
1	0	5	
Middle East and North Africa	Sub-Saharan Africa	United States & Canada	

2022-007-IG-MR

Case number

PC-10534

Public comment number

United States and Canada

Region

Indie

Commenter's first name

Pereira

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

Critique is not hate speech

Full Comment

Facebook seems to perceive any critique of a group of people or of a country or culture as hate speech. This is terribly simplistic and risks the real harms that come when evil deeds are allowed to flourish without critique. I too have been banned by Facebook for posting a critique of my own culture which was mislabeled as hate speech. I have also reported clear and obvious hate speech that wasn't removed. Facebook is an abject failure at discerning the difference.

Poetry is very frequently censored on Facebook whenever that poetry critiques harmful organizations or groups. This disappearing of honest critique by Facebook serves to aid oppressors and abusers while leaving their victims voiceless and hopeless that their cry for justice will be heard.

Link to Attachment

No Attachment

2022-007-IG-MR

Case number

PC-10618

Public comment number

Central and South Asia

Region

Shmyla Khan

Commenter's first name

& Noor Waheed

Commenter's last name

English

Commenter's preferred language

Digital Rights Foundation

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

This comment deals with the complexities of artistic expression, positing that the current policy regarding ‘Violence and Incitement’, particularly concerning coded statements, is inadequate to account for the nuances of such expression.

Furthermore, current transparency practices by Meta regarding law enforcement requests are completely inadequate, leaving users vulnerable to arbitrary decisions and state over-reach. Lastly, ‘accuracy at scale’ is both a policy and resource issue: policies need to be responsive to local contexts—not just at the national level, but at the local and community level—and implemented through greater resources allocated for content moderation that accounts for the local context.

Full Comment

Determining whether drill music can be posted on Meta’s platforms has wide-ranging implications for expression on the internet and the power asserted by Law Enforcement Agencies (LEAs) over content. This is a matter of particular interest as an organization based in Pakistan, where the state regularly silences expression under the pretext of “security”. Despite the crackdown by UK LEAs, drill music is undeniably a thriving subculture. Beyond its portrayal as a rallying call for gang-violence, it serves as a medium for disenfranchised youth, particularly black and brown, to express their discontent with the system that perpetuates discrimination and exclusion. Though some have connected this music with increased violence, more recent studies have demonstrated that the lyrics and imagery in drill music find no causal link with violence. The current war on drill music ignores the possibility of “adverse selection bias” where people who are more likely to commit violent crimes (as a product of socio-economic forces) happen to share an affinity with that music. While drill music does reference violence frequently, it is worth noting that much of the violence is symbolic, used to build the “reputation and

identity” of gangs. Any determination of art such as drill music cannot be understood without the context of overcriminalization of the communities that produce it. In the UK, two rappers have previously received a suspended one-year sentence for performing their song at a concert. In this backdrop, it is important for Meta to develop specific and clear standards based on international human rights law (IHRL). Artistic expression is protected under Articles 19 of the UDHR and ICCPR. These Articles notably cover cultural and artistic expression, even expression some might find offensive. Broadly, the genre of drill music falls under these categories. Under IHRL, exceptions exist for protection of national security and public order. These limitations must be (i) provided for by law, (ii) serving a legitimate aim (iii) necessary for a legitimate purpose. Meta’s Community Guidelines regarding ‘Violence and Incitement’ must fall within these exceptions. The policy regarding coded statements, which is likely to impact artistic expression that relies on inferences and symbolic speech, currently does not elaborate on the nature of veiled and implicit threats. It goes beyond a clear and present danger standard of imminent violence, can be read as covering generalized expressions of violence. The Guidelines as they stand could result in a widespread ban on drill rap and must be fine tuned to determine whether speech is part of a larger culture and expression of violence, or is directly resulting in violence? This question is crucial because over-regulation of art, particularly based on determinations made by LEAs, has the potential of stifling subcultures, curtailing social media’s ability to become avenues of expression for marginalized communities, and perpetuating racialized logics of policing on digital platforms. As part of its commitment to freedom of expression, Meta’s policies must factor in allowances for humor, satire, and artistic expression. Technology can be leveraged to introduce mechanisms to protect these categories of speech by allowing creators to pre-mark their content as “satire”. Further, to protect its users against widespread and imminent harm Meta can codify the time/manner/place restrictions mentioned earlier in its policies; they can temporarily remove the content and restore it once the threat has been resolved. In processing takedown requests by LEAs, a number of factors must be taken into account. Firstly, whether the takedown request gave substantial and specific reasons to link the content to incitement to violence. Meta should consider a high evidentiary threshold. Secondly, logics presented by LEAs in cases such as these must guard against “street illiteracy” which results in misunderstanding of content and art produced by marginalized groups. Thirdly, content moderators handling requests from LEAs must take into account historical and political factors determining the relationship of content producers with the state in order to contextualize the reasons by LEAs. Furthermore, it is also important to ensure the competence of content moderators and their ability to make complex determinations such as balancing public safety with artistic freedoms. Standards that allow for these determinations at scale should not preclude localized content moderation models that are responsive to the local context, agile enough to take specific complexities into account while implementing generalized Community Guidelines. For this particular case, extreme measures like “operation: domain” and “criminal behavior orders” issued by UK police to ban even the performance of drill music seem unnecessarily restrictive, and orders by LEAs to social media platforms

appear to be an extension of the same. This is supported by the fact that creating music with “violent” lyrics” and imagery is not against the law in the UK, and studies find no definitive or causal relationship between real-life crimes and drill music. The OSB, in taking its decision, would benefit from engaging with communities of drill music creators to better understand these dynamics. Lastly, given the outsized power of state bodies such as LEAs to impact speech and influence social media platforms, Meta should practice radical transparency for every request it receives from LEAs, documenting each request as it occurs: description of the offending content, the grounds of reporting content (including but not limited to what section of community guidelines was violated), the number of requests complied with and why, number of requests denied and in such a case on what grounds did Meta deny these requests.

Link to Attachment

[PC-10618](#)

2022-007-IG-MR

Case number

PC-10623

Public comment number

United States and Canada

Region

Nadine

Commenter's first name

Farid Johnson

Commenter's last name

English

Commenter's preferred language

PEN America

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

PEN America finds Meta's removal of the track "Secret's Not Safe" by drill rapper Chinx (OS) and the associated music video to be an inappropriate act of censorship. While we agree that Meta has a responsibility to address true incitements to violence on its social media platforms, we disagree that this type of artistic expression constitutes such an incitement.

Full Comment

To the members of the Oversight Board, Thank you for the opportunity to comment on case number 2022-007-IG-MR, UK drill music. PEN America, the United States based chapter of the PEN International movement, stands at the intersection of literature and human rights to protect free expression in the United States and around the globe. Our PEN Charter calls us to uphold "the principle of unhampered transmission of thought within each nation and between all nations." We champion the freedom to write and work to unite writers and their allies to celebrate creative expression and defend the liberties that make it possible. PEN America finds Meta's removal of the track "Secret's Not Safe" by drill rapper Chinx (OS) and the associated music video to be an inappropriate act of censorship. While we agree that Meta has a responsibility to address true incitements to violence on its social media platforms, we disagree that this type of artistic expression constitutes such an incitement. Meta has framed the crucial questions of this case as assessing "whether, in this case and more generally, the safety risks associated with the potential instigation of gang violence outweigh the value of artistic expression in drill music." This framing, however, largely assumes exactly that which should be denied: that drill or rap music, and associated music videos, constitute potential or actual instigations of violence. Meta's stated reason for removal of the song and video was their referencing of past specific violent incidents, which, combined with other violent imagery, was deemed sufficient to consider the music a veiled or

implicit threat. This inference does not appear to take into account either the context of the music nor its conventions, in terms of its origination in predominantly Black communities and its use as an outlet to voice frustrations about oppressive institutions. An understanding of drill music, or rap more generally, as involving “implied threats” ignores this well-known context. As such, actions taken to censor this type of music, or otherwise punish those who make it, with this mistaken understanding in mind do little more than stifle artistic expression. Mere references to violent incidents, or the inclusion of violent imagery, cannot be an adequate standard for considering something an incitement to violence. Many well-known poems, novels, and songs include such imagery while being widely and uncontroversially shared across social media platforms. When this fact is taken into account, actions to censor art forms like drill appear to be another way in which minority groups have their rights infringed on the basis of an unequal or inconsistent standard. Some have tried to draw more specific connections between drill music and gang violence in order to argue for its suppression or censorship. For example, this report claims that drill music “played a role” in 1-in-3 gang-related homicides in 2018 in the U.K., but that characterization was applied such that if anyone involved in the homicide (perpetrator or victim) was an “aspiring” drill musician, or drill music videos were used as evidence in the trial, then drill music “played a role.” This assertion neither establishes a causation between drill music and crime, nor recognizes the context of drill music’s rising popularity, which is largely in disadvantaged communities that were already affected by violence. As PEN America argued in this letter, which urged New York lawmakers to ban the use of artists’ work as evidence in court, context matters. The circumstances and artistic tradition within which a piece of expression is created are essential to determining whether the piece involves a credible threat. The use of drill music as evidence in trials does not show, without further information, that there is a connection between drill and rising violence. A piece of music’s use in a trial is legitimate only if there is already a reason to believe that there is a genuine connection between that music and the crime—but an assumption along these lines is necessarily circular. Drill music that includes “violent” imagery does not necessarily constitute an actual incitement to violence. Failure to acknowledge the circumstances of and the artistic tradition behind the creation of a piece of expression, leads to the type of situation exemplified by this case, in which artistic expression is taken out of context and improperly used as evidence of incitement. Meta’s policies must account for this kind of artistic context. This would better serve the cause of public safety, by not only clarifying what constitutes a credible threat, but also strengthening rights to free expression online. Lastly, the series of events leading to this content removal, including the request by law enforcement without an alleged illegality and the failure of Meta to inform the user of the reason for the initiation of removal, is troubling. At the very least, the platform should be as transparent as possible in cases where the government is launching a complaint, should have reported this to the user whose content was removed, and should incorporate incident information and numbers in its general content moderation and transparency reporting.

Link to Attachment

[PG-10623](#)

2022-007-IG-MR

Case number

PC-10629

Public comment number

Latin America and Caribbean

Region

Juanita

Commenter's first name

Castro

Commenter's last name

English

Commenter's preferred language

Karisma Foundation

Organization

Yes

Response on behalf of organization

Short summary provided by the commenter

This type of moderation in cases involving issues of freedom of artistic expression should be resolved by a team of experts within Meta in charge of the procedure, since it involves boundaries between rights that are very difficult to detect by automated processes of mass moderation. Likewise, the appeal processes must be improved so that the accused persons can know who is making the complaint and why, but also so that they can have the right to a fair and balanced defense.

Full Comment

Since its beginnings, urban music, particularly rap has been characterized by battles between performers of the genre, most of them remain only in the musical field, although some have transcended to the physical plane. This music is recognized worldwide for being the music chosen by young people from popular and marginal neighborhoods to express their discontent with society, to deal with ideological differences and to express their emotions. The drill is a subgenre of this urban music, characterized by having a higher degree of violent content in their songs, is related to music such as rap, reggaeton and trap, in recent years has gained great relevance among the hip hop scene in England.

We believe that these exceptional cases should not lead to a generalized censorship of this musical form of debate in the world of hip hop, one of the most important urban cultures in the world and a means of expression, mainly of non-conformist sectors of society. Attempting to moderate this online content should take into account: what is the purpose of the content, what is the past of these disputes and what is the probability that this discussion will leave the virtual world and incite acts of physical violence? To moderate all types of language or content that are linked to protest music, which by nature causes discomfort in some sectors of society, like most artistic expressions, would constitute an act of violation of the freedom of artistic expression on the digital public sphere of internet.

This special procedure must integrate a factual and legal study (documents), carried out by a team of experts. Meta already gives special treatment to government requests for user data , so this methodology can be applied to requests for content moderation -review or removal of content- by a government. Therefore, this type of request must be in writing, detailed and with a clear language. Moreover, the user account involved in the request must have a participation in this process. However, the content moderation request should be more demanding, since governments must not only allege the violation of the platform's rules but should also explain how it does not violate human rights standards. The latter means that governments must justify the law enforcement requests for the review or removal of content under Meta policies and also demonstrate that the Freedom of Expression limitation considers human rights parameters, such as the tripartite test. It is worth saying that this requirement is NOT an exorbitant burden for government requests due to the fact that they are already obliged to justify any limitation on freedom of expression under international law.

Currently, government requests for user data are included in the transparency report with data and graphs, however, it is necessary to incorporate clear and complete qualitative information. That is, Meta must have a public protocol explaining the step-by-step analysis of content moderation requests. In addition, all requests should be published in the transparency report, with a detailed explanation of the reasons why the content moderation requests by the government was accepted or not, including the contextual analysis of each law enforcement request by a government. And, it would be recommended that each analysis is presented publicly through Meta's transparency tool, in the language of the country involved in the government content moderation request.

Link to Attachment

[PC-10629](#)

2022-007-IG-MR

Case number

-

Commenter's first name

PC-10971

Public comment number

-

Commenter's last name

United States and Canada

Region

English

Commenter's preferred language

Electronic Frontier Foundation

Organization

Short summary provided by the commenter

See Attachment

Full Comment

See Attachment

Link to Attachment

[PC-10971](#)

Yes

Response on behalf of organization

