



Candidate for Mayor Assassinated in Mexico

2024-053-FB-MR, 2024-054-IG-MR, 2024-055-IG-UA, 2024-056-IG-UA

Summary

In four cases of videos showing the assassination of Mexican mayoral candidate José Alfredo Cabrera Barrientos, the Board notes how Meta treated posts differently when three of them should have benefited from the same outcome – to remain up under the newsworthiness allowance. These three posts were shared by news outlets clearly reporting on a political assassination ahead of Mexico’s elections: Meta left two up but removed one. Taking down reports on issues being debated by the public limits access to essential information and hinders free speech. This is concerning given the risks that news outlets face in Mexico when reporting on state corruption and organized crime. While there was an uneven application of the newsworthiness allowance in these cases, the Board also sets out its concerns about the effectiveness of the allowance itself. To address this, the Board reiterates its recent recommendation from the [Footage of Moscow Terrorist Attack](#) decision, calling for an exception to be made to the rule that does not allow third-party imagery showing the moment of designated attacks on visible victims. This updated approach would help ensure fairer treatment for all users.

About the Cases

In May 2024, four pieces of content about the assassination of a candidate running for mayor in the Mexican state of Guerrero were either posted by or reshared from news media accounts in Latin America. All four posts include similar videos showing José Alfredo Cabrera Barrientos on the campaign trail before a gun is aimed at him, followed by blurry images and sounds of gunshots.



The first two cases involve posts shared by large media organizations. The caption for the first post discusses how many candidates have been murdered during the election cycle, while the audio for the video includes a statement by the state prosecutor's office explaining that Cabrera Barrientos was under protection when killed. It was viewed about 59,000 times. The second post includes a warning about the video's sensitivity and a caption reporting on the Governor of Guerrero's statement condemning the murder. It was viewed more than a million times.

As Meta had designated this assassination as a violating violent event under its Dangerous Organizations and Individuals policy, another version of the video had already been added to a Media Matching Service (MMS) bank, which is programmed to remove the same content. Under the policy, users are not allowed to share third-party imagery depicting the moment of such designated attacks on visible victims. The first two posts, which were identified by the MMS bank and referred to Meta's subject matter experts for additional review, were left up despite breaking Meta's rules. They were given a newsworthiness allowance, occasionally granted for content Meta decides has high public interest value. Both posts remain on Meta's platforms with "Mark as Disturbing" warning screens and newsworthy labels, but were referred to the Board.

In the third and fourth cases, users appealed Meta's decisions to remove their posts to the Board. The third post involved a reshare of the video, with a message imposed on it stating that an uncensored version was available on Telegram. This had 17,000 views. The fourth post included a caption noting who had been shot and injured at the scene. It had 11,000 views. After an MMS bank identified both posts, they were removed.

The assassination of Cabrera Barrientos took place on the final day of campaigning ahead of Mexico's nationwide elections on June 2. Political violence has been a feature of recent elections in the country, with organized crime partially responsible. This has led candidates to drop out of election races, fearful for their lives.

Key Findings

While Meta was right to keep up the first two posts on its platforms as newsworthy content, the Board finds the company was not right to take down the post in the fourth



case from Instagram. This post also had high public interest value. There was no material difference to justify a different outcome. Even after the Board selected the fourth case, Meta failed to apply the same newsworthiness allowance, stating this post sensationalized the footage by informing users it had gone viral. However, this detail is included alongside other information about the shooting, including details about the number of casualties, the Governor’s statement and the fact the shooter was killed at the event. Although Cabrera Barrientos is visible and identifiable, he was a public figure attending an election rally, so the privacy concerns reduce and the public interest value outweighs risks of harm.

On the third post, which directed users to a Telegram link for an uncensored, graphic version of the video to get around Meta’s prohibition on sharing third-party imagery of attacks on visible victims, the majority of the Board agrees with Meta that this content posed greater risks to security and privacy – and should have been taken down. For the majority, Meta was right not to grant a newsworthiness allowance, especially given the post had no additional caption or commentary indicating its purpose was to inform others or condemn the assassination. A minority of the Board disagrees, finding that the third post should also qualify for the newsworthiness allowance, as it is similar to the others.

As the Board recently noted in its [Footage of Moscow Terrorist Attack](#) decision, imagery of designated attacks can be shared for multiple reasons. While Meta is concerned about such content glorifying, supporting or representing criminal groups’ activities, the rule that does not allow users to share third-party imagery of designated attacks on visible victims is leading to removal of content with low or no risk of harm. Of relevance to these cases, experts have noted that criminal groups in Mexico do not generally use videos of political assassinations for recruitment purposes, although they may share them to intimidate. Furthermore, the Board found no evidence of this footage having been recorded by the perpetrators or being used to inspire copycat behavior.

While the Board found the newsworthiness allowance should be applied to the fourth post, it notes that the allowance is rarely used since there are limited ways for Meta to identify content to benefit from it. In combination with the multiple factors that need



to be considered to grant the allowance, this increases the risks of the allowance’s random application, to the detriment of users. This is why the Board believes that a change to Meta’s policy, as highlighted in our Footage of Moscow Terrorist Attack decision, is preferable to Meta’s current approach.

The Oversight Board’s Decision

The Oversight Board upholds Meta’s decisions in the first three cases. The Board overturns Meta in the fourth case, requiring the post to be restored with a “Mark as Disturbing” warning screen.

The Board reiterates its recommendation from the recent Footage of Moscow Terrorist Attack decision, stating that Meta should allow, with a “Mark as Disturbing” warning screen, third-party imagery of a designated event showing the moment of attacks on visible but not personally identifiable victims for news reporting, awareness raising or to condemn.

*Case summaries provide an overview of cases and do not have precedential value.

Full Case Decision

1. Case Description and Background

On May 30, 2024, four different accounts posted about the assassination of José Alfredo Cabrera Barrientos, who was running for mayor in the municipality of Coyuca de Benitez in the Mexican state of Guerrero. He had been shot and killed the day before during a campaign rally. All four pieces of content, one on Facebook and three on Instagram, were either posted by or reshared from news media accounts based in Latin America. The posts include similar videos, showing Cabrera Barrientos shaking hands with constituents before a gun is aimed at him. Blurred or blurry images follow the sound of multiple gunshots



and people screaming. Each post is accompanied by a caption, in Spanish, providing facts about the shooting.

Meta designated the assassination of Cabrera Barrientos as a violating violent event under its Dangerous Organizations and Individuals policy. This means, among other things, that users are not allowed to share third-party imagery depicting the moment of such designated attacks on visible victims. Meta's subject matter experts had previously assessed another version of the video as violating and added it to a Media Matching Service (MMS) bank that was programmed to remove this content.

The first post was shared by a large media organization and includes a caption stating that 23 candidates for political office have been murdered during Mexico's current election cycle. The audio accompanying the footage provides more details, including a statement by the state prosecutor's office explaining the shooter had been killed at the event and the fact that Cabrera Barrientos was under protection when he was killed. The post was viewed about 59,000 times.

The second post, also shared by a large media organization, includes a warning added by the user that the video is sensitive. The caption reports on a statement by the Governor of Guerrero, in which she condemns the killing and expresses condolences to the family. It was viewed more than a million times.

These two posts were referred to the Board by Meta. After being identified by an MMS bank programmed to automatically remove this content, the posts were escalated to Meta's subject matter experts for additional review. The Board has previously described systems where this type of escalation might occur (see, for example, [Meta's Cross-Check Program policy advisory opinion](#)). The subject matter experts determined the posts violated the Dangerous Organizations and Individuals policy. However, they granted a newsworthiness allowance to keep the posts on the platform due to their public interest value. The posts therefore



remained on the platform with a “Mark as Disturbing” warning screen and newsworthy label.

In the third case, a user reshared content from a different media organization, without adding anything to it. The reshared video captures the candidate moments before his assassination, including the moment when a gun is aimed at him. The caption provides information about the assassination without any additional context. There is a message imposed on the video, which is restated in the caption, instructing viewers that an “uncensored” video is available on Telegram. It was viewed about 17,000 times.

The fourth post was shared by a media organization, with a caption noting that one of the attackers was shot at the scene and, in addition to the candidate, three others were injured. It was viewed about 11,000 times. These posts were removed after an MMS bank identified them. Both users in these two cases appealed to the Board.

The Board notes the following context in reaching its decision.

José Alfredo Cabrera Barrientos was the candidate for a coalition of the opposition political parties PRI-PAN-PRD, running for the position of Coyuca de Benitez’s Mayor. The assassination took place on the final day of campaigning ahead of nationwide elections on June 2, 2024. At the time, Cabrera Barrientos was under special [protection measures](#), with a security team in place. Reports, including the first and fourth posts in these cases, indicate that one attacker was shot and killed at the event. At least one other person suspected of being involved was taken into custody and later [found dead](#) while in detention.

This assassination took place within a wider context of political violence in Mexico. During the 2018 election cycle, organized crime was [reportedly](#) responsible for about half of the political violence, as “politicians or political



candidates are identified as rivals when they don't cooperate with criminal groups, which can turn them into targets for assassination or threats.”

During the 2021 election cycle, United Nations (UN) and regional human rights experts [reported](#) 250 political murders in the pre-electoral and campaigning period in Mexico. Those experts further noted “at least 782 other politically motivated attacks – ranging from death threats to attempted murder – against politicians.”

This violence has a chilling effect on candidates. [According to UN](#) human rights experts, in the 2021 electoral cycle “many candidates dropped out, citing fears for their lives.” International and regional experts further highlighted the impact this has on “the right of citizens to elect the candidate of their choice.” In the most recent 2024 cycle, [over 8,000 candidates](#) for office reportedly dropped out of their races, an increase from previous elections. The context of political violence was reported to be a contributing factor. The Inter-American Commission on Human Rights ([IACHR](#)) also condemned violence against candidates: “Since last year, [it] has observed with concern the occurrence of a series of acts of violence, including murders, threats, and kidnappings against pre-candidates, candidates, and leaders or activists of different political movements or affiliations.” According to the IACHR, from March 2024 to May 24, 2024, at least 15 pre-candidates or candidates were murdered, along with nine other individuals who had either expressed interest in running or were unofficial candidates.

According to the [Committee to Protect Journalists](#) (CPJ) and the [Global Initiative Against Transnational Organized Crime](#), Mexico is one of the most dangerous countries in the world for journalists. According to Freedom House's [2024 Report on Mexico](#): “Gangs have engaged in threats and violence against bloggers and online journalists who report on organized crime. Self-censorship has increased, with many newspapers in violent areas avoiding publishing stories concerning organized crime.” [Journalists](#) trying to report on the link between government



officials and criminal gangs have been murdered, leading to further silencing and fear.

Criminal gangs in Mexico are [reportedly](#) “active on Facebook... [and use the platform to intimidate] rival groups and civilians.” However, experts consulted by the Board indicate that criminal groups in Mexico do not generally use videos of political assassinations as a recruitment tool but may share violent imagery, to intimidate opponents, including journalists.

2. User Submissions

The news outlet that posted the content in the second case, which Meta kept up on Instagram as newsworthy content, submitted a statement to the Board. The submission states that information about the assassination is important to share, given the electoral context in Mexico. The post included important factual background about the assassination and reported on the statement released by the Governor of Guerrero. The users who posted the content in the third and fourth cases appealed Meta’s removal decisions to the Board. In their statements to the Board, they say they are reporting important news about violence and terrorism. Both express frustration that they have been censored.

3. Meta’s Content Policies and Submissions

1. Meta’s Content Policies

Dangerous Organizations and Individuals Community Standard

The [Dangerous Organizations and Individuals](#) policy rationale states that, in an effort to prevent and disrupt real-world harm, Meta does not allow organizations or individuals that proclaim a violent mission or are engaged in violence to have a presence on its platforms. The Community Standard also prohibits “content that glorifies, supports, or represents events that Meta designates as violating violent events,” including “terrorist attacks” and “multiple-victim violence or



attempted multiple-victim violence.” Meta prohibits: “(1) glorification, support or representation of the perpetrator(s) of such attacks; (2) perpetrator-generated content relating to such attacks; or (3) *third-party imagery depicting the moment of such attacks on visible victims*,” (emphasis added).

According to internal guidelines for reviewers, Meta removes imagery depicting the moment of attacks on visible victims “regardless of sharing context.” Meta does not require the victim to be visible at the same time as the violence, as long as it is clear the violence is directed at the victim, who is visible at some point in the footage.

Violent and Graphic Content Community Standard

The [Violent and Graphic Content](#) policy rationale states that the company understands people “have different sensitivities with regard to graphic and violent imagery.” Meta therefore removes the most graphic content while allowing and adding a warning label to other graphic content. This policy allows, with a “Mark as Disturbing” warning screen, “imagery (both videos and still images) depicting a person’s violent death (including their moment of death or the aftermath) or a person experiencing a life threatening event.” The warning screen limits visibility to users aged over 18 and doesn’t recommend the content to users who do not follow the account. The policy prohibits such imagery when it shows dismemberment, visible innards, burning or throat slitting.

Newsworthiness Allowance

In certain circumstances, the company will allow content that may violate its policies to remain on the platform if it is “[newsworthy](#) and if keeping it visible is in the public interest.” When making the determination, “[Meta will] assess whether that content surfaces an imminent threat to public health or safety, or gives voice to perspectives currently being debated as part of a political process.” According to Meta, its analysis is informed by country-specific circumstances, the nature of the speech and the political structure of the country affected. Meta can also apply a warning screen to content that it keeps up under



this allowance and limit users under 18 from viewing the content. Lastly, the company states: “Newsworthy allowance can be ‘narrow,’ in which an allowance applies to a single piece of content, or ‘scaled,’ which may apply more broadly to something like a phrase.”

II. Meta’s Submissions

Meta designated the assassination of José Alfredo Cabrera Barrientos as a violating violent event under its Dangerous Organizations and Individuals policy soon after the attack. Meta determined that all four posts violated the company’s policy prohibiting “third-party imagery depicting the moment of [designated] attacks on visible victims.”

The company explained that it generally removes all such designated imagery, regardless of the context in which it is shared, for two main reasons. The first reason is safety. According to the company, removing this content helps to limit copycat behaviors (imitative behaviors) and avoid the spread of content that raises the profile of and may have propaganda value to the perpetrators. The second reason is the privacy and dignity of victims and their families. The company also aims to protect the dignity of any victims and their loved ones “who did not consent to being the subject of public curiosity and media attention.”

Designating certain attacks allows Meta to quickly remove content under its Dangerous Organizations and Individuals policy across its platforms in response to key events.

Meta stated that it grants few newsworthiness allowances for content that violates this policy. According to Meta, given the concerns its policy aims to address, these allowances are typically narrow in scope and generally limited to footage shared by recognized media outlets for news reporting.



For the first two cases, the company issued a newsworthiness allowance considering the “wide national reach” of the two news outlets that posted the content, and the fact that the footage was contextualized with captions. The company assessed the posts as having high public interest value, due to the relevance of the violence and insecurity associated with the election cycle to the public debate. Meta did not rule out all risks, due to “the proximity to election day, particularly of copycat attacks against other candidates in areas that lack security, as well as a potential risk of dignitary harm to the family of the candidate.” Nevertheless, the company considered the fact that the news organizations took “editorial steps to avoid sharing imagery in a sensationalist way,” and “included captions that contextualized the footage within the broader context of how the violence and insecurity have impacted the electoral cycle and shared information on the official law enforcement response to the incident.” The first post does not include the exact moment of the shooting and the second post includes its own warning screen.

When Meta granted a “narrow” newsworthiness allowance for these two posts, the company applied a label (or “newsworthy inform treatment”) to let users know the posts were allowed for the purpose of public awareness. It also applied a “Mark as Disturbing” warning screen, which prevented users under 18 from viewing the content. Any user who reshared these two specific posts also benefited from the allowance. No posts by other accounts received a newsworthiness allowance related to footage of the assassination.

All other content identified by the Media Matching Service (MMS) bank as violating footage of the Cabrera Barrientos shooting was automatically removed from Meta’s platforms. Meta configured the MMS bank to remove content without applying a strike, “to ensure enforcement was proportional given the possibility users could be sharing the footage to raise awareness about or condemn the attack.”



The third and fourth posts were therefore removed by the MMS bank without a strike being applied to the users' accounts. Once the Board brought these two posts to Meta's attention, the company confirmed they did not merit a newsworthiness allowance. The third and fourth posts were not shared by "well-known news outlets, nor did they contextualize the video in the same way" as the first and second posts. Meta took note of the fact that the third post directed users to "uncensored images" on Telegram and that the fourth post emphasized that the imagery had gone viral on social media. The company found this sensationalized the footage.

According to Meta, the company's decisions to remove the third and fourth posts were in line with the conditions of legality, legitimacy, and necessity and proportionality. First, Meta reiterated the necessity of generally removing imagery showing the moment of attack of designated events, given the risks that the content may promote copycat behavior and advance the aims of perpetrators. Removing the content, but not applying a strike to these users, was the least restrictive means of addressing the risk of harm.

The Board asked questions on how many newsworthiness allowances the company issued for imagery of the assassination, the status of accounts and pathways for users to benefit from the newsworthiness allowance, whether Meta had specialized teams in place to address heightened risks during the election cycle and how these teams were prepared. Meta responded to all questions.

4. Public Comments

The Oversight Board received 10 public comments that met [the terms for submission](#). Seven of the comments were submitted from Latin America and the Caribbean, two from the United States and Canada, and one from Europe. To read public comments submitted with consent to publish, click [here](#).



The submissions covered the following themes: electoral and political violence during Mexico’s 2024 general election; the impact of political violence on democratic processes; how Meta should moderate content and adjust its policy on third-party sharing of violating violent events imagery; the effectiveness of the newsworthiness allowance; the role of social media in providing information about election processes; the use of social media by criminal organizations; general information about standards in Mexico for depicting political violence in news reporting, and the importance of freedom of speech in the context of elections in Mexico.

5. Oversight Board Analysis

The Board selected these cases to address how political violence is depicted on Meta’s platforms and its potential impact on electoral processes. These cases fall within the [Board’s strategic priority](#) of Elections and Civic Space.

The Board analyzed Meta’s decisions in these cases against Meta’s content policies, values and human rights responsibilities. The Board also assessed the implications of these cases for Meta’s broader approach to content governance.

5.1 Compliance with Meta’s Content Policies

1. Content Rules

All four posts violate Meta’s prohibition on “third-party imagery depicting the moment of [designated] attacks on visible victims.” Meta designated the assassination of José Alfredo Cabrera Barrientos immediately after the event on May 29, 2024. All four posts include footage showing Cabrera Barrientos moving through the crowd, as well as the moment the gun is pointed at him and immediately afterwards, to the sound of gunshots and people screaming. The rule, under the Dangerous Organizations and Individuals Community Standard, and further explained in Meta’s internal guidelines, prohibits such footage regardless of the context in which it is shared.



Meta was right to keep the first two posts on its platforms as newsworthy content, applying a “Mark as Disturbing” warning screen and a newsworthy label. Under its policies, Meta should also have allowed the fourth post to remain on Instagram due to its public interest value. There was no material difference between these posts to justify a different outcome. After the Board selected this case, despite the content being reviewed by subject matter experts entitled to determine newsworthiness allowances (or other measures only applied on escalation), Meta still failed to correct this differential treatment. This goes against the principle of treating users fairly.

The content in these three posts shows a shooting at a campaign event in an election cycle during which political violence was a central issue. The first, second and fourth posts provided information about the shooting, including the number of casualties and statements released by the Governor in response. The Board disagrees with Meta that the fourth post sensationalizes the footage by informing users it had gone viral on social media. Rather than sensationalizing, this highlights the post’s significance to the public. That information is included along with other relevant details on the number of casualties, including that the shooter was killed at the event, and the statement released by the Governor of Guerrero.

When journalists limit their coverage of key events, such as the killing of a politician, public access to critical information is limited. Given the significant risks that news outlets and journalists face in Mexico, ensuring the accessibility of this type of news on online platforms is vital, especially during an election period. Consequently, the Board considers threats against journalists, and the resulting self-censorship, as relevant context for its newsworthiness analysis.

Additionally, while the victim is fully visible and identifiable in the footage, the fact he was a public figure lessens the privacy concerns in this case. He was attending a public campaign rally during an election, and not depicted in a



humiliating or degrading manner. For these three posts, the public interest value outweighs the risks of harm.

For the third post, the majority of the Board agrees with Meta that the content poses greater risks and it finds Meta’s decision not to grant a newsworthiness allowance was reasonable. The content provides a video of the assassination without any additional information or caption that suggests an intent to report, raise awareness or condemn the attack.

On the contrary, the message imposed on the video, and restated in the caption, informs viewers that an “uncensored” version of the video is available on Telegram and provides a link to this platform. The majority of the Board finds that the post aims explicitly to circumvent the prohibition on sharing third-party imagery of attacks on visible victims by directing users to violating content on an external platform. For these reasons, the majority agrees that Meta was right not to grant a newsworthiness allowance in this case. Additionally, in its research, the Board also verified that the linked Telegram channel highlights extremely violent footage, including imagery of beheadings and, relevant to this case, graphic imagery of the candidate’s assassination.

For a minority of Board Members, the third post, being similar to the others, also deserves the newsworthiness allowance. Insofar as the decision of the majority relied on the fact that this post included a hyperlink to another platform, the minority believes that a hyperlink, by itself, should not be seen as “publication” of the content to which it refers.

When Meta granted a newsworthiness allowance for the first two posts, the company added a newsworthy label to inform users that the posts were allowed for public awareness. The Board has previously recommended that Meta notify users when content remains on the platform due to a newsworthiness allowance (see [Colombia Protests](#) decision, recommendation no. 4, and [Sudan Graphic Video](#) decision, recommendation no. 4). The Board welcomes this practice, as it



provides people with valuable context for why policy-violating content is allowed to stay on the platform.

5.2 Compliance with Meta’s Human Rights Responsibilities

The Board finds that keeping the first two posts up, with a warning screen and a newsworthy label, and removing the third post was consistent with Meta’s human rights responsibilities. However, the Board finds removing the fourth post was not consistent with Meta’s human rights responsibilities.

Freedom of Expression (Article 19 ICCPR)

Meta’s content moderation practices can have adverse impacts on the right to freedom of expression. Article 19 of the [International Covenant on Civil and Political Rights](#) (ICCPR) provides broad protection for this right, given its importance to political discourse, and the Human Rights Committee has noted that it also protects expression that may be considered “deeply offensive,” ([General Comment No. 34](#), paras. 11, 13 and 38). Article 19’s protection is “particularly high” for “public debate in a democratic society concerning figures in the public and political domain,” ([General Comment No. 34](#), para. 34).

When restrictions on expression are imposed by a state, they must meet the requirements of legality, legitimate aim, and necessity and proportionality (Article 19, para 3, ICCPR). These requirements are often referred to as the “three-part test.” The Board uses this framework to interpret Meta’s voluntary human rights commitments, in relation both to the individual content decisions under review and to Meta’s broader approach to content governance. As the [UN Special Rapporteur on freedom of opinion and expression](#) has stated, although “companies do not have the obligations of Governments, their impact is of a sort that requires them to assess the same kind of questions about protecting their users’ right to freedom of expression,” ([A/74/486](#), para. 41).

I. Legality (Clarity and Accessibility of the Rules)



The principle of legality requires rules limiting expression to be accessible and clear, formulated with sufficient precision to enable an individual to regulate their conduct accordingly ([General Comment No. 34](#), para. 25). Additionally, these rules “may not confer unfettered discretion for the restriction of freedom of expression on those charged with [their] execution” and must “provide sufficient guidance to those charged with their execution to enable them to ascertain what sorts of expression are properly restricted and what sorts are not,” ([General Comment No. 34](#), para. 25). The UN Special Rapporteur on freedom of expression has stated that when applied to private actors’ governance of online speech, rules should be clear and specific ([A/HRC/38/35](#), para. 46). People using Meta’s platforms should be able to access and understand the rules, and content reviewers should have clear guidance regarding their enforcement.

The Board previously discussed and recommended how Meta could better structure its rules around designated events in its [Footage of Moscow Terrorist Attack](#) decision. Here, the Board reiterates that while Meta should improve this policy, its rule prohibiting third-party footage of designated events on visible victims is sufficiently clear for users to understand that content like this is prohibited. The footage shared in these posts depicts a shooting that targeted the candidate and resulted in multiple victims. The policy provides sufficiently clear notice to users that this kind of footage can be designated.

II. Legitimate Aim

Meta’s Dangerous Organizations and Individuals policy aims to “prevent and disrupt real-world harm.” In several decisions, the Board has found that this policy pursues the legitimate aim of protecting the rights of others, such as the right to life ([ICCPR](#), Article 6) and the right to non-discrimination and equality ([ICCPR](#), Articles 2 and 26) because it covers organizations that promote hate, violence and discrimination as well as designated violent events motivated by hate. See [Referring to Designated Dangerous Individuals as “Shaheed,” Sudan’s](#)



[Rapid Support Forces Video Captive, Hostages Kidnapped from Israel](#) and [Greek 2023 Elections Campaign](#) decisions. Meta’s policies also pursue the legitimate aim of protecting the right to privacy (ICCPR, Article 17) of identifiable victims and their families (see [Video After Nigeria Church Attack](#) decision).

III. Necessity and Proportionality

Under ICCPR Article 19(3), necessity and proportionality requires that restrictions on expression “must be appropriate to achieve their protective function; they must be the least intrusive instrument amongst those which might achieve their protective function; they must be proportionate to the interest to be protected,” (General Comment No. 34, para. 34).

The Board recognizes that, in developing the designation policy, Meta has erred on the side of safety and privacy. Meta explained that its current policy approach allows the company to swiftly remove this content through MMS banks, which helps disrupt the spread of perpetrator propaganda and can limit copycat behavior. Removing this content also helps protect the privacy and dignity of victims and their families when victims are visible. As the Board recently noted in the *Footage of Moscow Terrorist Attack* decision, a narrower rule risks underenforcement of content depicting violent events. It could also allow footage to be reused for harmful purposes that Meta may struggle to detect and remove. In some contexts, the risks of incitement or repurposing of such imagery do justify erring on the side of safety.

However, the Board also emphasized in its *Footage of Moscow Terrorist Attack* decision that imagery of designated attacks can serve multiple purposes. Not all content depicting a designated attack, as in three of these cases, serves to glorify, support or represent criminal groups’ activities. Such content does not always have the outcomes Meta aims to prevent. Policies that prioritize overenforcement, regardless of context, pose risks to freedom of expression, access to information and public participation. This rule does lead to removal of content with low or no risk of harm.



To address potential overenforcement, Meta has several policy tools. Three of those tools are especially pertinent here, though others exist as well. First, it can remove content without applying strikes or other penalties that might restrict the user. Withholding strikes on content enforced by MMS banks mitigates the risks of limiting access for users through feature limits or account suspension and serves as an important tool for ensuring proportionality.

Second, Meta can also apply newsworthiness allowances to permit designated content with limited potential to create risks to public safety and the dignity of those depicted. However, for the newsworthiness allowance to be an effective mitigation measure on overenforcement, it must be effectively applied to relevant content. In previous cases and the policy advisory opinion on cross-check, the Board has identified multiple obstacles to the effectiveness of the allowance (see [Meta’s Cross-Check Program policy advisory opinion](#), [Sudan’s Rapid Support Forces Video Captive](#), [Armenian Prisoners of War Video](#)). The newsworthiness allowance can only be applied on escalation and not by at-scale moderators. Because Meta’s at-scale moderators are not instructed or empowered to identify and escalate content that could benefit from the newsworthiness allowance, there are limited pathways for Meta to identify content it should consider for the newsworthiness allowance. For news outlets, journalists and others reporting on public interest issues not enrolled in Meta’s cross-check program or with access to Meta’s internal teams, it will be difficult to gain access to those within the company empowered to consider and apply the newsworthiness allowance. Additionally, the decision to grant the allowance requires considering multiple factors to balance public interest and potential harm, leading to a lack of predictability and increasing the risk of arbitrariness in its application, to the detriment of users.

The effect is that the allowance is rarely used (see [Sudan Graphic Video](#) decision). From June 1, 2023 through June 1, 2024, Meta has [reported](#) 32 allowances. In these cases, for example, two posts were identified and escalated while one was not, despite similarity in content and context, undermining the fair treatment of users.



The Board finds that in these specific cases, given the context in Mexico, it was neither necessary nor proportionate to remove the first, second and fourth posts.

The first, second and fourth posts do not contain elements suggesting risks of recruitment or incitement to copycat behavior. Experts consulted by the Board stated that criminal groups in Mexico do not generally use videos of political assassinations for recruitment, but may share such content to intimidate. The Board found no evidence that the footage in these cases was recorded by the perpetrator nor that the shooter or criminal groups shared these specific posts to inspire copycat behavior, spread perpetrator’s propaganda or glorify their violent acts.

On the contrary, these three posts were shared by news outlets reporting on a political assassination at a campaign rally days before an upcoming election. Removing reports on issues being debated and scrutinized by the public, such as violence and the state’s response, would limit access to essential information and hinder free speech, while providing marginal gains in safety. In its Footage of Moscow Terrorist Attack decision, the Board noted that images of attacks often evoke stronger reactions than abstract descriptions. Images humanize victims and elicit moral outrage, sympathy, awareness of violence and encourage accountability. Given the significant risks journalists and news outlets face in Mexico when reporting on state corruption and organized crime, limiting their access to social media is especially concerning. Additionally, as the victim was a public figure engaged in public acts and he was not depicted in a humiliating or degrading manner, there are more limited privacy interests.

In these three cases, applying a “Mark as Disturbing” warning screen, under Meta’s Violent and Graphic Content Community Standard, is a less restrictive means to protect the rights to safety and privacy. When Meta applies a warning screen, several consequences follow. All users must click through a screen to view content, and it is not available to users under the age of 18. Furthermore, the content is then removed from recommendations to users who do not follow



the account (see [Al-Shifa Hospital](#) and [Hostages Kidnapped From Israel](#) decisions). These measures ensure that child users are not exposed to the content and limit the reach of this content to users who have sought it out.

For the third post, the majority of the Board considers that the content presents greater risks to security and privacy. In that post, the user reshared content from a media account that included a message directing viewers to an “uncensored” video on Telegram. The majority agrees with Meta that removing the post is necessary and proportionate to protect safety. By sharing the post with a link to view graphic imagery of an individual’s death with no additional caption or commentary, the user gave no clear indication that their purpose was to inform others or to condemn the violence. Lacking such indications, and linking to uncensored footage, the post clearly suggests that the user was aiming to circumvent Meta’s Community Standards regarding Dangerous Organizations and Individuals. A minority of Board Members disagree, asserting that removing the third post was neither necessary nor proportionate.

On the proportionality of Meta’s response, the Board welcomes the fact that the company did not apply strikes against the users who posted the two pieces of content that were removed and determined that in some circumstances, there is no need for additional penalization in the form of a strike. The Board emphasizes the value of separating Meta’s enforcement actions on content from the penalties given to users. It also recognizes that withholding strikes constitutes an important tool for achieving proportionality (see [Iranian Make-up Video for a Child Marriage](#) decision), insofar as the requirement of proportionality takes into account the restriction’s imposition not only on the interests of others, including listeners, but also on the interests of the speaker (UN Special Rapporteur on freedom of expression, Special Rapporteur Communication [No. USA 6/2017](#), pg. 3).

The Board analyzed these cases in accordance with the newsworthiness allowance, as it reflects Meta’s current policy approach to these posts. However, as previously mentioned, the newsworthiness allowance has multiple limitations in accessibility and predictability. Public comments highlighted similar concerns, as well as the fear that users may self-censor to avoid account



level penalties ([PC-30727](#) Digital Speech Lab). For these reasons, the Board considers that the newsworthiness allowance is not the most effective or least restrictive approach available to Meta.

The Board recently highlighted these same concerns on the prohibition of third-party imagery depicting the moment of designated attacks on visible victims in the Footage of Moscow Terrorist Attack decision. That decision concluded that the most effective way to protect freedom of expression while mitigating harm and the risk of copycat behavior would be to establish an exception within the policy. This exception would permit third-party imagery of a designated event depicting the moment of attacks on visible victims, when shared in the contexts of news reporting, condemnation or awareness-raising. The content would have a “Mark as Disturbing” warning screen. Meta is currently assessing this recommendation.

The Board reiterates that its proposed approach would better respect rights. To meet Meta’s safety concerns, the company could also require that users posting content for news reporting, condemnation or awareness-raising make their intent clear, as it does under the Dangerous Organizations and Individuals policy. The Board notes Meta defines awareness-raising in its guidance as “sharing, discussing or reporting new information ... for the purpose of improving the understanding of an issue or knowledge of a subject that has public interest value. Awareness raising ... should not aim to incite violence or spread hate or misinformation,” (see [Reporting on Pakistan Parliament Speech](#) and [Communal Violence in Indian State of Odisha](#) decisions). The company could continue to remove unclear or ambiguous content, deferring to safety concerns. Meta could also choose to apply the exception only on-escalation, if clear protocols for identifying content are provided. While the Board consistently expresses concerns about the effectiveness of escalations-only policies (see [Sudan’s Rapid Support Forces Video Captive](#) and [Sudan Graphic Video](#) decisions), it believes that a clearly articulated and policy-specific exception enforced on-escalation is preferable to relying on the newsworthiness allowance (see [Armenian Prisoners of War Video](#) decision).



Under the framework proposed by the Footage of Moscow Terrorist Attack decision, the same outcome would be reached here without the need for the application of the newsworthiness allowance. The first, second and fourth posts would remain on platform as news reporting. Given that the intent behind the third post was not to report, raise awareness or condemn, the third post should be removed. While the Footage of Moscow Terrorist Attack decision addressed third-party footage with visible but not personally identifiable victims, the victim in this case is identifiable. However, given that he is a public figure at a public event, and he is not depicted in a humiliating or degrading manner, the privacy interests involved are similarly reduced and the content should benefit from the recommended exception.

By limiting reliance on the rarely granted and unpredictable newsworthiness allowance, a clear policy exception for news reporting, condemnation and awareness-raising would help Meta to treat users fairly.

6. The Oversight Board’s Decision

The Oversight Board upholds Meta’s decisions to leave up the first and second posts, and to remove the third post.

The Oversight Board overturns Meta’s decision to take down the fourth post, requiring it to be restored with a “Mark as Disturbing” warning screen.

7. Recommendations

Content Policy

The Oversight Board reiterates its previous recommendation in the Footage of Moscow Terrorist Attack decision:

Meta should allow, with a “Mark as Disturbing” warning screen, third-party imagery of a designated event showing the moment of attacks on visible but not personally identifiable victims when shared in news reporting,



condemnation and awareness-raising contexts (Footage of Moscow Terrorist Attack decision, recommendation no. 1).

***Procedural Note:**

- The Oversight Board’s decisions are made by panels of five Members and approved by a majority vote of the full Board. Board decisions do not necessarily represent the views of all Members.
- Under its [Charter](#), the Oversight Board may review appeals from users whose content Meta removed, appeals from users who reported content that Meta left up, and decisions that Meta refers to it (Charter Article 2, Section 1). The Board has binding authority to uphold or overturn Meta’s content decisions (Charter Article 3, Section 5; Charter Article 4). The Board may issue non-binding recommendations that Meta is required to respond to (Charter Article 3, Section 4; Article 4). When Meta commits to act on recommendations, the Board monitors their implementation.
- For this case decision, independent research was commissioned on behalf of the Board. The Board was assisted by Duco Advisors, an advisory firm focusing on the intersection of geopolitics, trust and safety, and technology.