

Videos of Teachers Hitting Children 2025-020-FB-MR, 2025-021-FB-MR

Summary

In considering two videos showing non-sexual child abuse in educational settings, the Oversight Board recommends an exception be included in Meta's Child Sexual Exploitation, Abuse and Nudity Community Standard to allow such content, when shared to condemn, report or raise awareness, and when children are not identifiable. Meta's current prohibition on posting videos or photos of non-sexual child abuse, regardless of intent and even when identities are obscured, results in disproportionate restrictions on freedom of expression. Sharing such content would contribute to public debate on important children's rights issues.

The Board has upheld Meta's decision to allow this content in one case and overturned Meta's decision to take it down in the other. Both posts should stay up under a newsworthiness allowance, with warning screens.

About the Cases

In the first case, an Indian media organization posted a video on its Facebook page in which a teacher yelled at a young school student for not studying, repeatedly hit his head and back, and appeared to pull at his turban. A superimposed blurred patch covered the child's face for most of the video. The caption noted a state official has called for accountability.

In the second case, a video was posted on a Facebook page that appears to share local news from a region in France. The video shows a group of very young children, with one child crying. A teacher hits the child, who falls to the ground while the other children watch. All faces in the video are blurred. The caption references the school, date and neighborhood where the incident took place and an investigation.



After it was reported and escalated, Meta's policy experts determined the Indian content violated the Child Sexual Exploitation, Abuse and Nudity policy and removed it. As the content was posted to raise awareness, Meta did not apply a strike against the account, the company revealed to the Board.

The French content was removed without human review for violating the same policy. After the content creator appealed, Meta confirmed the takedown was correct but it removed the strikes it had applied due to public interest and because the content was shared to raise awareness.

Meta referred both cases to the Board. When preparing its submissions, Meta's policy experts decided to keep up the French content with a newsworthiness allowance and warning screen. Meta said an attorney representing the child's parents had shared the video in local media to raise awareness of the incident. For Meta, this meant the public interest value outweighed the harm, as the "parents' consent mitigated the privacy and dignity concerns."

Key Findings

The Board finds that Meta's prohibition on posting videos or photos of non-sexual child abuse, regardless of intent and even when identities are obscured, results in disproportionate restrictions on freedom of expression. The Child Sexual Exploitation, Abuse and Nudity policy does not differentiate between identifiable and non-identifiable children. However, identifiability is critical. If this risk is mitigated, there are more limited privacy and dignity concerns.

The Board finds both posts violate the Child Sexual Exploitation, Abuse and Nudity policy. However, a majority finds they should remain up under a newsworthiness allowance, with warning screens and visibility restricted to users over 18-years-old. This is consistent with Meta's human rights responsibilities and better meets the test of necessity and proportionality, allowing greater freedom of expression. Both posts have



high public interest value, encourage accountability and try to prevent identification of the children. In the French case, parental consent, while not unassailable, supports the presumption that the video's dissemination is not contrary to the child's best interests. Such reporting contributes to important public debates on children's rights, in the context of growing global efforts to ban child abuse in educational settings.

A minority of the Board disagrees, finding the public interest did not outweigh the risk of harm to the privacy and dignity of the children depicted. For these members, removal would best respect the interests of the children.

The Oversight Board's Decision

The Board upholds Meta's decision in the French case to keep the content on platform with a newsworthiness allowance and a warning screen. The Board overturns Meta's decision in the Indian case and requires the content to be restored with a newsworthiness allowance and a warning screen.

The Board recommends that Meta:

- Include an exception in its public-facing Child Sexual Exploitation, Abuse and Nudity policy allowing images and videos of non-sexual child abuse perpetrated by adults, when shared with the intent to condemn, report and raise awareness. This must only be applied when the child is neither directly identifiable by name or image nor functionally identifiable (when contextual clues are likely to lead to the identification of the individual). Content should be allowed with a warning screen and restricted visibility to users aged 18 and older. This exception should be applied on escalation only.
- Should not apply strikes to accounts whose non-sexual child abuse content is removed on escalation where there are clear indicators of the user's intent to condemn, report or raise awareness.



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

Full Case Decision

1. Case Description and Background

This decision addresses two cases of videos posted on Facebook showing non-sexual child abuse in educational settings. The videos were posted to raise awareness. Meta determined both pieces of content violated its Child Sexual Exploitation, Abuse and Nudity policy prohibiting "videos or photos that depict non-sexual child abuse regardless of sharing intent." Meta referred both cases to the Board for guidance on addressing the safety and dignity of children, as well as the need to raise awareness about newsworthy events and issues.

In the first case, a media organization in India posted a video on its Facebook page in which a teacher yelled at a young school student for not studying. She repeatedly hit his head and back, and appeared to pull at his turban. The face of the child was superimposed with a blurred patch. Although occasionally the patch failed to cover his moving face, it remained difficult to clearly identify him. The teacher and other students were visible. The caption identified where the incident took place and noted a state official had called for accountability.

The post was viewed several thousand times. Ten people reported it. Because the account receives cross-check protections, one of the reports was escalated to policy experts, who determined the content violated the Child Sexual Exploitation, Abuse and Nudity policy and removed the post. Meta describes cross-check as a mistake-prevention strategy that provides for extra levels of review. Meta revealed to the Board that it did not apply a strike against the content creator's account because the content was posted to raise awareness.



In the second case, a video was posted on a Facebook page that appears to share local news from a region in France. The video shows a group of very young children in an educational setting, with one child crying. The teacher hits the child, who falls to the ground while the other children watch. The teacher also appears to spray something on the child. All faces are blurred in this video. The caption cites the specific neighborhood, date and school where the incident was apparently filmed and references an investigation.

The post was viewed several thousand times. A user reported it and an automated system identified it as potentially violating the Child Sexual Exploitation, Abuse and Nudity policy. The content was then removed without human review for violating the policy. Meta applied a standard and a severe strike to the case content creator's account. The administrator of the page escalated the content through a "creator support channel" to appeal the removal. Meta maintains resources for content creators who seek to grow their audience and earn money, and may include additional support channels for users. Policy experts then confirmed the removal was correct but reversed the earlier decision to apply strikes because "of the public interest and awareness raising context."

Subsequently, when Meta was preparing its submissions to the Board, its policy experts decided to allow the content on the platform with <u>a newsworthiness allowance</u> and warning screen. According to Meta, an attorney representing the child's parents had shared the video in local media to raise awareness of the incident. For the company, this meant the public interest value outweighed the harm, as the "parents' consent mitigated the privacy and dignity concerns."

The Board notes the following context in reaching its decision:

The United Nations Educational, Scientific and Cultural Organization (<u>UNESCO</u>) reports that school violence is widespread, affecting students and education staff. It estimates that one billion children aged two to 17 face some form of such violence each year. According to the <u>World Health Organization</u> (WHO), "In some countries, nearly all



students report being physically punished by school staff," with the <u>highest rates</u> observed in Africa and South Asia. A <u>UNICEF</u> – the United Nations Children's Fund - study found that in South Asia, corporal punishment and bullying in schools remain common despite legislative bans.

According to the civil society organization <u>End Corporal Punishment</u>, corporal punishment is banned in schools in India. The law only applies to children aged six to 14 and excludes certain settings, such as religious schools, day care and the home. In contrast, <u>France has banned</u> corporal punishment in all settings, including the home, since 2019.

2. User Submissions

Following Meta's referrals of these cases and the Board's decision to accept them, both users who posted the content were provided with an opportunity to submit a statement. No response was received.

3. Meta's Content Policies and Submissions

I. Meta's Content Policies

Child Sexual Exploitation, Abuse and Nudity

Meta's policy rationale states that the company does "not allow content or activity that sexually exploits or endangers children." The Community Standard prohibits "videos or photos that depict real or non-real non-sexual child abuse regardless of sharing intent, unless the imagery is from real-world art, cartoons, movies or video games." The policy also prohibits "content that praises, supports, promotes, advocates for, provides instructions for or encourages participation in non-sexual child abuse." Meta shared with the Board that its internal guidance uses an "exhaustive list of specific acts of non-sexual physical abuse conducted by an adult or an animal towards anyone under the age of 18" to define non-sexual child abuse.



There are a small number of context-specific provisions in the policy that, according to Meta, are applied upon escalation to specialized teams. The current policy allows "videos or photos of non-sexual child abuse" to remain on platform on escalation when requested by law enforcement, child protection agencies or trusted safety partners, specifically to aid in efforts to bring a child to safety. It also allows "videos or photos that depict police officers or military personnel committing non-sexual child abuse." In both cases, Meta applies a disturbing content warning screen and restricts visibility to users aged 18 and older. The <u>Additional Protection of Minors policy</u> also states that Meta complies with "government requests for removal of non-sexual child abuse imagery."

There is no policy-level exception for depictions of non-sexual child abuse when shared as news reporting, awareness-raising or condemnation. However, the company may allow this content via its general newsworthiness allowance, after escalated review by policy teams (see Images of Partially Nude Indigenous Women).

Newsworthiness Allowance

In certain circumstances, the company will allow content that violates its policies to remain on platform if it is "newsworthy and if keeping it visible is in the public interest." Meta states that when making this determination it balances the public interest in the content against the risk of harm. The company assesses "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." The company says its analysis is informed by country-specific circumstances, the nature of the speech and the relevant political structure, and the degree of press freedom. For content that may be sensitive or disturbing that Meta keeps on platform via this allowance, the company includes a warning screen. It can also restrict access to users aged 18 and older. The allowance can be applied across any content policy, but because only specialized moderators may grant it, it happens very rarely. In its Transparency Center, Meta states that it applied the allowance 32 times between June 2023 and June 2024.



II. Meta's Submissions

Meta "takes a firm stance against sharing non-sexual child abuse content, regardless of intent, to prioritize the safety, dignity, and privacy of the minor," due to the "significant risk of harm" such content poses. According to Meta, a wide range of stakeholders highlighted that sharing this material can retraumatize the child, expose them to ongoing harassment and shame, and hinder their recovery, especially given the enduring visibility of content on social media.

Meta emphasized that child rights advocates support prioritizing the safety and privacy of child victims, while recognizing that social media can play a positive role in raising awareness. Meta also noted that "privacy risks are diminished, but not eliminated, when the victim's face is not visible or blurred." To balance these concerns, the policy bans videos and photos of non-sexual child abuse but permits textual descriptions, which are often used in news reporting.

Meta determined that both posts violated its policy prohibiting depictions of "real or non-real non-sexual child abuse regardless of sharing intent." However, Meta ultimately found the content in the French case was eligible for a newsworthiness allowance and restored the post with a warning screen. The allowance was granted due to the public interest value of the post and the limited risks. The child's parents consented to the video's distribution, which mitigated privacy and dignity concerns. Meta also noted that the child's face had been blurred. Meta stated that if either of these factors had been different, the outcome might have changed. Meta also acknowledged that blurring a child's face reduces but does not eliminate risk, as other contextual clues may lead to identification.

In contrast, the post in the Indian case was removed and did not receive a newsworthiness allowance. Meta determined that, although the content had public interest value to raise awareness about child abuse in schools, the risk of harm was significant. The post disclosed the school location and showed the teacher's face without blurring, further reducing anonymity. As a result, local individuals could



potentially identify the child. Meta also noted that, unlike the French case, there was no evidence that the child or their family had consented to the imagery being shared. Therefore, the public interest did not outweigh the risk of harm.

Meta shared with the Board that it did not ultimately apply a strike to either content creator, as both posts were shared to raise awareness. The French post received a standard and severe strike after the first review, but this decision was reversed shortly thereafter during escalated review.

The Board asked Meta questions about its non-sexual child abuse policy, including the rationale behind the policy, the stakeholders consulted and their views, the application of the newsworthiness allowance and the policy enforcement processes. The Board also inquired about the feasibility of providing users with tools to make children unidentifiable. Meta responded to all questions.

4. Public Comments

The Board received seven public comments that met <u>the terms for submission</u>. Six of the comments were submitted from the United States and Canada, and one was from Europe. To read public comments submitted with consent to publish, click <u>here</u>.

The submissions covered the following themes: the potential harms to children, their families and society from allowing non-sexual child abuse imagery on social media; social media platforms' responsibilities; corporal punishment as a violation of children's human rights; and reporting on child abuse as a matter of public interest, for accountability.

5. Oversight Board Analysis

The Board selected these cases given the importance of freedom of expression that supports accountability around matters of public interest and the need to respect the



best interests of the child. These cases highlight the tension between Meta's values of voice and ensuring the privacy, safety and dignity of children.

The Board analyzed Meta's decision in this case against the company's content policies, values and human rights responsibilities. The Board also assessed the implications of this case for Meta's broader approach to content governance.

5.1 Compliance With Meta's Content Policies

Content Rules

The Board finds that both posts violate Meta's Child Sexual Exploitation, Abuse and Nudity policy prohibiting depictions of non-sexual child abuse. Both videos show adults hitting children.

A majority of the Board finds that Meta was right to apply a newsworthiness allowance in the French case and that Meta should have also applied the newsworthiness allowance to the Indian case. For the majority, under Meta's policies, both pieces of content should remain on platform with warning screens and visibility should be restricted to users over 18-years-old.

In accordance with Meta's newsworthiness allowance test, both pieces of content had high public interest value. The videos depict acts of violence by adults, in positions of power and educational responsibility, against children, and each has captions that reflect efforts for accountability for this abuse. The majority emphasized the particular importance that visual depictions can have in journalism, advocacy and public awareness.

Both pieces of content also present limited risk to the children depicted. The majority highlighted efforts to obscure the identification of the children by blurring their faces. While information about location in the captions and contextual clues in the content could facilitate identification of the schools, the majority finds it would still be difficult



to identify the blurred children. Although Meta highlighted a risk of harm in the Indian case, the majority finds this harm unclear given identification is unlikely. Additionally, for the French case, the consent of the abused child's parents, while not dispositive, supports a presumption that the dissemination of the video is not contrary to the child's best interests.

A minority of the Board disagrees, finding that the newsworthiness allowance should not have been applied to either case. These Board Members agree there is public interest in the content but find that it does not outweigh the risk of harm to the privacy and dignity of the children depicted. The minority focused on the permanence of content once it has been disseminated online and the possibility of identifying the child victims, despite the blurring, through contextual clues in both posts. For the minority, parental consent does not diminish those risks. These Board Members believe that even if some parents consent to sharing such videos on social media, it does not necessarily serve the best interests of the child or children involved.

5.2 Compliance With Meta's Human Rights Responsibilities

A majority of the Board finds that keeping the content up with a warning screen and visibility restricted to people over 18-years-old in both cases is consistent with Meta's human rights responsibilities. Meta's removal of the post in the Indian case was not necessary or proportionate. A minority of the Board disagrees, finding that, in line with Meta's current policies, removal is necessary to protect the rights of the children depicted in the videos.

Freedom of Expression (Article 19 ICCPR)

On March 16, 2021, Meta announced its <u>Corporate Human Rights Policy</u> in which it outlines its commitment to respecting rights in accordance with the <u>United Nations Guiding Principles on Business and Human Rights</u> (UNGPs). According to the UNGPs, companies should "avoid infringing on the human rights of others and should address adverse human rights impacts with which they are involved" (Principle 11, UNGPs).



Article 19 of the <u>International Covenant on Civil and Political Rights</u> (ICCPR) provides broad protection for freedom of expression and the United Nations (UN) Human Rights Committee has noted that it protects commentary on public affairs, discussion of human rights, and journalism (<u>General Comment No. 34</u>, para. 11). The committee has also stated that expression is a necessary condition for ensuring accountability, which is essential for the promotion and protection of human rights (<u>General Comment No. 34</u>, para 2).

Any state restrictions on expression must comply with 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 human rights responsibilities in line with the UNGPs, both for specific posts and Meta's approach to content governance. As noted by the UN Special Rapporteur on freedom of expression, while companies do not have the same obligations as governments, their impact is of a sort that requires them to assess the same considerations regarding users' right to expression (A/74/486, para. 41).

Article 3 of the UN <u>Convention on the Rights of the Child</u> (UNCRC) states that a child's best interests must be a primary consideration in all actions concerning them. The Committee on the Rights of the Child (CRC) describes the best interest of the child as a flexible principle that applies to children of all ages, adapted to their specific circumstances and evolving development (<u>General Comment No. 14</u>, pp. 5-6). In line with this, <u>UNICEF principles and guidelines</u> for media reporting on children emphasize that children's rights and dignity must be respected in all situations, and that their best interests should take precedence over any other consideration, including advocacy and the promotion of child rights.

Article 19 of the UNCRC requires states to protect children from all forms of physical violence "while in the care of parent(s), legal guardian(s) or any other person who has the care of the child." The CRC has made it clear that all forms of violence against children are unacceptable, further noting that "corporal punishment is invariably



degrading" and "incompatible with the Convention" (General Comment No. 8, paras. 7, 11, 12, and General Comment No. 13, paras. 17, 24). Corporal punishment violates the child's dignity and exceeds acceptable school discipline limits (Article 28.2, UNCRC, General Comment No. 1, para. 8). Children must be protected from violence and harm in all environments, including in the digital environment (General Comment No. 25, para. 82).

Article 16 of the UNCRC guarantees children's right to privacy. The CRC further highlights that "privacy is vital to children's agency, dignity and safety and for the exercise of their rights" and that "threats may arise from ... a stranger sharing information about a child" online (General Comment No. 25, para. 67).

The Board's analysis is therefore informed by both strong protections for freedom of expression and the rights of the child, in particular their right to be protected from violence and their right to privacy.

I. Legality (Clarity and Accessibility of the Rules)

The principle of legality requires that any restriction on freedom of expression follows an established rule, which is accessible and clear to users (General Comment No. 34, para 25). The Child Sexual Exploitation, Abuse and Nudity policy prohibits "videos or photos that depict real or non-real non-sexual child abuse regardless of sharing intent." The Board finds that the rule prohibiting depictions of non-sexual child abuse is sufficiently clear as applied to these cases.

While the Board welcomes that Meta has provided greater clarity about the process and criteria for the newsworthiness allowance in its Transparency Center in response to Board recommendations, the Board reiterates its concerns about its application. The allowance has limited predictability and accessibility, due to the lack of clear pathways for identifying content that may qualify (see, among others, <u>Candidate for Mayor Assassinated in Mexico</u> and <u>Images of Partially Nude Indigenous Women</u>).



Any restriction on freedom of expression should also pursue one or more of the legitimate aims listed in the ICCPR, which include protecting the rights of others. Meta stated its rule prohibiting depictions of non-sexual child abuse "furthered the 'legitimate aim' of protecting the rights of others, including the right to privacy and protecting the dignity of the [children] depicted in the videos," in alignment with the best interests of the child. The Board, reiterating previous decisions assessing this standard, finds the policy serves the legitimate aim of protecting children's rights to physical and mental health (Article 19, UNCRC) and their right to privacy (Article 17, ICCPR; Article 16, UNCRC), consistent with respecting the best interests of the child (Article 3, UNCRC) (see Swedish Journalist Reporting Sexual Violence against Minors and News Documentary on Child Abuse in Pakistan).

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).

A majority of the Board disagrees with Meta's initial decisions to remove the French and Indian posts, finding that the limitation on expression was not necessary or proportionate to protect the rights of the child. The protection of freedom of expression in these cases is consistent with the protection of the rights of the child insofar as it furthers the aims of public awareness and accountability for violence against children in schools. The majority also finds that Meta's revised decision to allow the French post on platform is more consistent with its responsibilities to protect freedom of expression. Keeping up the content with a warning screen and permitting visibility only to users over 18-years-old better meets the tests of necessity and proportionality and allows greater freedom of expression. Lastly, the majority finds that the



newsworthiness allowance, with the more proportionate restrictions of a warning screen and age limitations, should also have been applied to the Indian post.

Human rights protections for this type of expression are high. Both pieces of content were shared by pages that are either media organizations or appear to share local news. The posts aimed to report and raise awareness about corporal punishment by teachers in school settings. This is a matter of high public interest and debate, particularly in India and South Asia. The posts included captions highlighting official efforts toward accountability. The majority considers that such reporting contributes to public debate on children's rights issues and helps expose abuse that might otherwise remain hidden.

The majority agrees that depictions of non-sexual child abuse may raise privacy and safety concerns for children. In these two cases, the faces of the child victims were blurred. They were not identified by name or image, and blurring limits the possibility of what Meta calls functional identification through contextual clues. Meta currently defines functional identification in this policy as identification "through means other than name or image if content includes information that is likely to lead to the identification of the individual." While the posts contain information about the location of the schools, the majority finds this information is not "likely to lead to the identification" of the specific individuals. In addition, in the French case, the majority notes that the factor of parental consent (together with the efforts to obscure the child's identity) supports the conclusion that allowing the video on Facebook is not contrary to the child's best interests. Arguments for removal based on the potential retraumatization of children from videos such as these are considerably weaker where meaningful steps to obscure the children's identity have been taken. This is particularly relevant when the content aims to raise public awareness and promote accountability for violence in schools, which ultimately advance the broader objective of protecting children from harm. The analysis of necessity and proportionality would, however, be quite different if the children could be easily identified. Identification is a key consideration here.



A minority of the Board disagrees. Given the potential harm to the depicted children's privacy, dignity and safety, these Members find removal would be the best way to respect the best interests of the child, as outlined in the UNCRC. They find that removal is both necessary and proportionate to that aim under Article 19(3). The minority, noting UNICEF's position, highlights that depictions of non-sexual child abuse on social media can re-traumatize child victims, and lead to public humiliation, stigmatization, bullying and exploitation (PC-31244). The minority emphasizes that these risks are amplified by the lasting nature of online content, particularly once it has gone viral, and the possibility of functional identification, despite blurring. For the minority, parental consent does not change this assessment, as online material may remain accessible indefinitely. By the time these children can voice their views it would be too late to mitigate the impact.

The minority also noted the restriction on expression is limited to imagery, and accountability and reform efforts can be pursued through other forms of expression. Experts consulted by the Board emphasized that reporting, narrative advocacy and legal petitions can effect change without focusing on imagery. This echoes UNICEF's view that accountability efforts must not come at the cost of a child's right to dignity and integrity.

Both the majority and minority agree that Meta was right not to apply strikes in this case. Meta's <u>Transparency Center</u> notes that it may or may not apply a strike to accounts posting content it removes, depending on several factors, including severity and context. The Board agrees that applying any penalty that impacts the ability of users to share additional content or use the platform would disproportionately impact their expression.

At the policy level, the Board finds that Meta's prohibition on posting videos or photos of non-sexual child abuse, regardless of intent and even when identities are obscured, results in disproportionate restrictions on freedom of expression. The Board notes that Meta's current policy does not differentiate between identifiable and non-identifiable children, focusing solely on whether the content depicts non-sexual child abuse.



However, identifiability is a critical factor. If this risk is mitigated, there are more limited privacy and dignity concerns.

International human rights standards provide strong protections for expression on matters of public interest and where it supports the realization of other rights. This includes efforts to report on incidents of child abuse in educational settings, particularly where violence against children in schools remains prevalent. This is especially relevant in the context of growing global efforts to ban such practices. Sharing this type of content can uncover systemic abuse, prompt public debate and lead to accountability or institutional reform, consistent with the best interests of the child.

Imagery also often evokes stronger reactions than narrative descriptions, as it provides vivid and compelling evidence and encourages accountability. In some regions, such as South Asia, the circulation of non-sexual child abuse imagery has led to immediate responses, including rescues, arrests or disciplinary actions.

The Board also notes that, according to expert analysis it commissioned, in most cases where non-sexual child abuse content is shared on social media, the intent is to raise awareness, report incidents, condemn abuse, support journalism or demand accountability. This pattern was confirmed in an internal Board study that reviewed other instances in which both videos were shared on different social media platforms and remain publicly available.

To better protect expression, while still respecting the best interests of the child, the Board recommends that Meta include an exception in its Child Sexual Exploitation, Abuse and Nudity policy to allow content depicting non-sexual abuse with two criteria. First, it should only apply when the child is neither directly identifiable by name or image nor functionally identifiable. Second, Meta should apply this exception when content is posted for news reporting, condemnation or awareness-raising.



Users should make their intent clear and Meta may continue to remove content that is ambiguous or unclear. This would further limit the risk of harm to children. Other Community Standards require users to clearly indicate their intent when creating or sharing otherwise prohibited content for the purposes of condemnation, reporting or raising awareness (for example, under the Dangerous Organizations and Individuals, Adult Nudity and Sexual Activity, Bullying and Harassment and Hateful Conduct policies).

The Board notes that Meta's Child Sexual Exploitation, Abuse and Nudity policy currently includes an exception for imagery of non-sexual abuse when committed by police officers or military personnel. That policy line contains no requirement of limited identifiability. The Board finds that non-sexual abuse committed by teachers engages similar accountability concerns. Yet, it believes limitations on identifiability are necessary to respect the privacy and safety of children, in line with the best interests of the child.

Meta should not apply strikes against users who share imagery of non-sexual child abuse with the intent to raise awareness, condemn or report news, even if the child is identifiable. This would be to better ensure the proportionality of its enforcement actions, which restrict freedom of expression. The Board understands that its recommended policy exception, particularly the need to consider the likelihood of functional identification, will be applied on escalation. While content with identifiable children should still be removed, Meta should not apply account penalties against users when there are clear indicators of their intent to report, raise awareness or condemn.

The Board acknowledges expert views and public comments that blurring or obscuring faces can help reduce harm but does not eliminate it entirely. Technical limitations, such as inconsistent blur strength and the presence of contextual clues (e.g., school uniforms, recognizable locations, references to age or other identifiers), may still allow for identification, posing risks to both the child and their family. Human rights standards further call for anonymity for child victims (UNCRC, General Comment No. 25, para. 57).



Blurring or obscuring a child's face remains a useful harm-reduction measure that lowers the likelihood of identification. In the News Documentary on Child Abuse in Pakistan decision, the Board encouraged Meta to ease user burden and reduce risks to children by providing users with more specific instructions and in-product tools like face-blurring for videos. Although technically feasible, Meta currently does not offer such tools (see Sudan's Rapid Support Forces Video Captive) and noted that development would require resources and consideration of legal and operational factors. https://www.oversightboard.com/decision/fb-14uy7pvn/ Given the importance of avoiding child identification, the Board again encourages Meta to develop tools to allow users to more effectively blur parts of content they wish to post, particularly if blurring would prevent a policy violation (see Sharing Private Residential Information, recommendation no. 13).

Content allowed under the police officer or military personnel exception receives a "mark as disturbing" warning screen and access is restricted to those over 18-years-old. These measures require users to click through to view the content and exclude it from being recommended to non-followers, as found in previous decisions (see Al-Shifa Hospital, Hostages Kidnapped From Israel and Candidate for Mayor Assassinated in Mexico).

The same measures should be taken for content in this new exception. In line with the legitimate aim of protecting the privacy and dignity of depicted children, warning screens and age restrictions are more proportionate than removal when the child is not identifiable. While the majority concluded that removal is not proportionate in these cases, given that blurring makes identification of the specific children unlikely, it agrees with the minority that some risk remains. Even if a child's face has been blurred or efforts to obscure children's faces have been taken, there still may be potential risks for their privacy or dignity. Because this risk is minimal, a less intrusive restriction on freedom of expression is justified. UNICEF highlighted the importance of limiting unintended exposure to such content, and that depicted children may face bullying and



stigmatization (<u>PC-31244</u>). Warning screens and age restrictions help address these concerns, limiting exposure and access, particularly among a child's peers.

6. The Oversight Board's Decision

The Oversight Board upholds Meta's decision in the French case to keep the content on platform with a newsworthiness allowance and a "mark as disturbing" warning screen. The Board overturns Meta's decision in the Indian case and requires the content to be restored to the platform with a newsworthiness allowance and the same warning screen.

7. Recommendations

A. Content Policy

1. To allow users to condemn, report and raise awareness of non-sexual child abuse, Meta should include an exception in its public-facing Child Sexual Exploitation, Abuse and Nudity Community Standard allowing images and videos of non-sexual child abuse perpetrated by adults, when shared with this intent. Content should be allowed with a "mark as disturbing" warning screen and restricted visibility to users aged 18 and older. In these cases, children must neither be directly identifiable (by name or image) nor functionally identifiable (when contextual clues are likely to lead to the identification of the individual). This exception should be applied on escalation only.

The Board will consider this recommendation implemented when Meta updates the public-facing Child Sexual Exploitation, Abuse and Nudity Community Standard in accordance with the above.

B. Enforcement



2. To ensure proportionate and consistent enforcement, Meta should not apply strikes to accounts whose non-sexual child abuse content it removes on escalation, where there are clear indicators of the user's intent to condemn, report or raise awareness.

The Board will consider this recommendation implemented when Meta shares its Internal Implementation Standards that incorporate this guidance for content reviewed on escalation.

*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 <u>Charter</u>, 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). Where 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.