



Holocaust Denial

2023-022-IG-UA

Summary

The Oversight Board has overturned Meta’s original decision to leave up an Instagram post containing false and distorted claims about the Holocaust. The Board finds that the content violated Meta’s Hate Speech Community Standard, which bans Holocaust denial. This prohibition is consistent with Meta’s human-rights responsibilities. The Board is concerned about Meta’s failure to remove this content and has questions about the effectiveness of the company’s enforcement. The Board recommends Meta take steps to ensure it is systematically measuring the accuracy of its enforcement of Holocaust denial content, at a more granular level.

About the Case

On September 8, 2020, an Instagram user posted a meme of Squidward – a cartoon character from the television series SpongeBob SquarePants. This includes a speech bubble entitled “Fun Facts About The Holocaust,” which contains false and distorted claims about the Holocaust. The claims, in English, question the number of victims of the Holocaust, suggesting it is not possible that six million Jewish people could have been murdered based on supposed population numbers the user quotes for before and after the Second World War. The post also questions the existence of crematoria at Auschwitz by claiming the chimneys were built after the war, and that world leaders at the time did not acknowledge the Holocaust in their memoirs.

On October 12, 2020, several weeks after the content was posted, Meta revised its Hate Speech Community Standard to explicitly prohibit Holocaust denial or distortion.

Since the content was posted in September 2020, users reported it six times for violating Meta’s Hate Speech policy. Four of these reports were reviewed by Meta’s automated systems that either assessed the content as non-violating or automatically closed the reports due to the company’s COVID-19 automation policies. These policies, introduced at the beginning of the pandemic in 2020, automatically closed certain review jobs to reduce the volume of reports being sent to human reviewers, while keeping open potentially “high-risk” reports.



Two of the six reports from users led to human reviewers assessing the content as non-violating. A user who reported the post in May 2023, after Meta announced it would no longer allow Holocaust denial, appealed the company’s decision to leave the content up. However, this was also automatically closed due to Meta’s COVID-19 automation policies, which were still in force in May 2023. They then appealed to the Oversight Board.

Key Findings

The Board finds that this content violates Meta’s Hate Speech Community Standard, which prohibits Holocaust denial on Facebook and Instagram. Experts consulted by the Board confirmed that all the post’s claims about the Holocaust were either blatantly untrue or misrepresented historical facts. The Board finds that Meta’s policy banning Holocaust denial is consistent with its human-rights responsibilities. Additionally, the Board is concerned that Meta did not remove this content even after the company changed its policies to explicitly prohibit Holocaust denial, despite human and automated reviews.

As part of this decision, the Board commissioned an assessment of Holocaust denial content on Meta’s platforms, which revealed use of the Squidward meme format to spread various types of antisemitic narratives. While the assessment showed a marked decline since October 2020 in content using terms like “Holohoax,” it found that there are gaps in Meta’s removal of Holocaust denial content. The assessment showed that content denying the Holocaust can still be found on Meta’s platforms, potentially because some users try to evade enforcement in alternative ways, such as by replacing vowels in words with symbols, or creating implicit narratives about Holocaust denial using memes and cartoons.

It is important to understand Holocaust denial as an element of antisemitism, which is discriminatory in its consequences.

The Board has questions about the effectiveness and accuracy of Meta’s moderation systems in removing Holocaust denial content from its platforms. Meta’s human reviewers are not provided the opportunity to label enforcement data in a granular way (i.e., violating content is labelled as “hate speech” rather than “Holocaust denial”). Based on insight gained from questions posed to Meta in this and previous cases, the Board understands these challenges are technically surmountable, if resource intensive. Meta should build systems to label enforcement data at a more granular level, especially in view of the real-world consequences of Holocaust denial. This would potentially improve its accuracy in moderating content that denies the Holocaust by providing better training materials for classifiers and human reviewers. As Meta increases its reliance on artificial intelligence to moderate content, the Board is interested in how the development of such systems can be shaped to prioritize more accurate enforcement of hate speech at a granular policy level.



The Board is also concerned that, as of May 2023, Meta was still applying its COVID-19 automation policies. In response to questions from the Board, Meta revealed that it automatically closed the user’s appeal against its decision to leave this content on Instagram in May 2023, more than three years after the pandemic began and shortly after both the World Health Organization and United States declared that COVID-19 was no longer a “public health emergency of international concern.” There was a pressing need for Meta to prioritize the removal of hate speech and it is concerning that measures introduced as a pandemic contingency could endure long after circumstances reasonably justified them.

The Oversight Board’s Decision

The Oversight Board overturns Meta’s original decision to leave up the content.

The Board recommends that Meta:

- Take technical steps to ensure that it is sufficiently and systematically measuring the accuracy of its enforcement of Holocaust denial content, to include gathering more granular details.
- Publicly confirm whether it has fully ended all COVID-19 automation policies put in place during the pandemic.

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

Full Case Decision

1. Decision Summary

The Oversight Board overturns Meta’s original decision to leave up an Instagram post including false and distorted claims about the Holocaust. The Board finds that the content violated Meta’s [Hate Speech](#) Community Standard, which prohibits Holocaust denial. After the Board selected the case for review, Meta determined that its original decision to leave up the content was in error and removed the post.

2. Case Description and Background

On September 8, 2020, an Instagram user posted a meme of [Squidward](#) – a cartoon character from the television series, *SpongeBob SquarePants* – which includes a speech



bubble entitled “Fun Facts About The Holocaust,” containing false and distorted claims about the Holocaust. The post, in English, calls into question the number of victims of the Holocaust, suggesting it is not possible that six million Jewish people could have been murdered based on supposed population numbers the user quotes for before and after the Second World War. The post also questions the existence of crematoria at Auschwitz by claiming that the chimneys were built after the war, and claims that world leaders at the time did not acknowledge the Holocaust in their memoirs.

The caption below the image includes several tags relating to memes, some of which target specific geographical audiences. The user who posted the content had about 10,000 followers and was not considered a public figure by Meta. In comments on their own post responding to criticism from others, the user reiterated that the false claims were “real history.” The post was viewed under 500 times and had fewer than 100 likes.

On October 12, 2020, several weeks after the content was originally posted, Meta [announced](#) revisions to its Hate Speech Community Standard to explicitly prohibit Holocaust denial or distortion, noting that “organizations that study trends in hate speech are reporting increases in online attacks against many groups worldwide,” and that their decision was “supported by the well-documented rise in anti-Semitism globally and the alarming level of ignorance about the Holocaust.” Meta added “denying or distorting information about the Holocaust” to its list of “designated dehumanizing comparisons, generalizations, or behavioural statements” within the Community Standard (Tier 1). Two years later, on November 23, 2022, Meta reorganized the Hate Speech Community Standard to remove the word “distortion” and list “Holocaust denial” under Tier 1 as an example of prohibited “harmful stereotypes historically linked to intimidation, exclusion, or violence on the basis of a protected characteristic.”

Since the content was posted in September 2020, users reported it six times for hate speech. Four of these reports were made before Meta’s October 12, 2020 policy change and two came after. Of the six reports, four were reviewed by automation and were either assessed as non-violating or auto-closed due to Meta’s “COVID-19 automation policies,” with the post left up on Instagram. According to Meta, its COVID-19 automation policies, introduced at the beginning of the pandemic in 2020, “auto-closed review jobs based on a



variety of criteria” to reduce the volume of reports being sent to human reviewers, while keeping open potentially “high-risk” reports.

Two of the six reports led to human reviewers assessing the content as non-violating, one prior to the October 2020 policy change and one after, in May 2023. In both instances, the reviewers determined the content did not violate Meta’s content policies and they did not remove the post from Instagram. The user who reported the content in May 2023 appealed Meta’s decision to leave the content up, but that appeal was also auto-closed due to Meta’s COVID-19-related automation policies, which were still in force at the time. The same user then appealed to the Board, noting in their submission that it was “quite frankly shocking that this [content] is allowed to remain up.”

The Board notes the following background in relation to antisemitism and Holocaust denial in reaching its decision in this case. In January 2022, the UN General Assembly adopted by consensus resolution [76/250](#), which reaffirmed the importance of remembering the nearly six million victims of the Holocaust and expressed concern at the spread of Holocaust denial on online platforms. It also noted the concerns of the UN Special Rapporteur on contemporary forms of racism (report [A/74/253](#)) that the frequency of antisemitic incidents appears to be increasing in magnitude in several regions, especially in North America and Europe. The resolution emphasizes that Holocaust denial is a form of antisemitism, and explains that “Holocaust denial refers specifically to any attempt to claim that the Holocaust did not take place, and may include publicly denying or calling into doubt the use of principal mechanisms of destruction (such as gas chambers, mass shooting, starvation, and torture) or the intentionality of the genocide of the Jewish people.”

The UN Special Rapporteur on freedom of religion or belief also emphasized in 2019 the growing use of antisemitic tropes, including “slogans, images, stereotypes and conspiracy theories meant to incite and justify hostility, discrimination and violence against Jews” (report [A/74/358](#), at para. 30.) The Board notes that Holocaust denial and distortion are forms of conspiracy theory and reinforce harmful antisemitic stereotypes, in particular the dangerous idea that Jewish people invented the Holocaust as fiction to advance purported plans of world domination.



Organizations such as the [Anti-Defamation League \(ADL\)](#) and the [American Jewish Committee](#) have reported a sharp increase in antisemitic incidents. The ADL researches and documents antisemitic content online, most recently in its [2022 Online Holocaust Denial Report Card](#) and two [August 2023 investigations](#) into antisemitic content on major platforms. It gave Meta a score of C on the report card, based on an alphabetical grading scale of A to F, with A being the highest score. In its [investigations](#), ADL pointed out that “Facebook and Instagram, in fact, continue hosting some hate groups that parent company Meta has previously banned as ‘dangerous organizations’.” The investigation also emphasized that the problem was particularly bad on Instagram, with the platform having “recommended accounts spreading the most virulent and graphic antisemitism identified in the study to a 14-year-old persona” created for the investigation.

3. Oversight Board Authority and Scope

The Board has authority to review Meta’s decision following an appeal from the person who previously reported the content that was left up (Charter Article 2, Section 1; Bylaws Article 3, Section 1).

The Board may uphold or overturn Meta’s decision (Charter Article 3, Section 5), and this decision is binding on the company (Charter Article 4). Meta must also assess the feasibility of applying its decision in respect of identical content with parallel context (Charter Article 4). The Board’s decisions may include non-binding recommendations that Meta must respond to (Charter Article 3, Section 4; Article 4). When Meta commits to act on recommendations, the Board monitors their implementation.

When the Board selects a case like this one, in which Meta subsequently acknowledges that it made an error, the Board reviews the original decision to increase understanding of the content moderation process, and to make recommendations to reduce errors and increase fairness for people who use Facebook and Instagram.

4. Sources of Authority and Guidance

The following standards and precedents informed the Board’s analysis in this case:

I. Oversight Board Decisions



- [Mention of the Taliban in News Reporting](#)
- [South Africa Slurs](#)
- [“Two Buttons” Meme](#)
- [Depiction of Zwarte Piet](#)
- [Armenians in Azerbaijan](#)
- [Removal of COVID-19 Misinformation](#)

II. *Meta’s Content Policies*

The Instagram [Community Guidelines](#) state: “It’s never OK to encourage violence or attack anyone based on their race, ethnicity, national origin, sex, gender, gender identity, sexual orientation, religious affiliation, disabilities or diseases. When hate speech is being shared to challenge it or to raise awareness, we may allow it. In those instances, we ask that you express your intent clearly.” Instagram’s Community Guidelines direct users to Facebook’s [Hate Speech Community Standard](#), which states that hate speech is not allowed on the platform “because it creates an environment of intimidation and exclusion and, in some cases, may promote real-world violence.”

Facebook’s Hate Speech Community Standard defines hate speech as a direct attack against people on the basis of protected characteristics, including race, ethnicity and/or national origin, and describes three tiers of attack. When the content was posted in September 2020, the Community Standards did not explicitly prohibit Holocaust denial in any tier. Tier 1 did, however, prohibit: “Mocking the concept, events or victims of hate crimes even if no real person is depicted in an image.” In response to questions from the Board, Meta explained that its Internal Implementation Standards currently list the Holocaust as a specific example of what it considers a “hate crime,” but that it does not keep logs of the changes to Implementation Standards and Known Questions in the same way it logs changes to Community Standards in the Transparency Center.

On October 12, 2020, Meta [announced](#) it was updating its Hate Speech policy “to prohibit any content that denies or distorts the Holocaust,” citing “the well-documented rise in anti-Semitism globally and the alarming level of ignorance about the Holocaust, especially among young people.” It also cited a recent survey of adults in the United States aged between 18 and 39, which showed that “almost a quarter said they believed the Holocaust



was a myth, that it had been exaggerated or they weren't sure." On the same day, Tier 1 of the policy was updated, adding "Denying or distorting information about the Holocaust" to a list of 10 other examples of "[d]esignated dehumanizing comparisons, generalization or behavioral statements (in written or visual form)."

On November 23, 2022, Meta updated its policy. It now prohibits content targeting a person or group of people based on protected characteristic(s) with "dehumanizing speech or imagery in the form of comparisons, generalizations or unqualified behavioral statements (in written or visual form) to or about: [...] Harmful stereotypes historically linked to intimidation, exclusion, or violence on the basis of a protected characteristic, such as [...] Holocaust denial."

The Board's analysis of the content policies was also informed by Meta's commitment to [voice](#), which the company describes as "paramount," and its values of safety and dignity.

III. Meta's Human-Rights Responsibilities

The UN Guiding Principles on Business and Human Rights (UNGPs), endorsed by the UN Human Rights Council in 2011, establish a voluntary framework for the human-rights responsibilities of private businesses. In 2021, Meta [announced](#) its [Corporate Human Rights Policy](#), in which it reaffirmed its commitment to respecting human rights in accordance with the UNGPs.

The Board's analysis of Meta's human-rights responsibilities in this case was informed by the following international standards:

- The rights to freedom of opinion and expression: Article 19, International Covenant on Civil and Political Rights ([ICCPR](#)), Article 20, para. 2 ICCPR, [General Comment 34](#), Human Rights Committee, 2011; UN Special Rapporteur (UNSR) on freedom of opinion and expression, reports: [A/HRC/38/35](#) (2018) and [A/74/486](#) (2019); International Convention on the Elimination of All Forms of Racial Discrimination ([ICERD](#)), Article 4(a).
- Equality and non-discrimination: Article 1, [Universal Declaration of Human Rights](#); Article 2, para. 1 and Article 26, ICCPR; UN General Assembly [Resolution 76/250](#) on Holocaust denial (2022).



5. User Submissions

The Board received two submissions from users in this case. The first was from the person who reported the content, with their submission forming part of their appeal to the Board. The second submission was from the person who posted the content, who was invited to submit a comment after the Board selected this case, following Meta taking action to reverse its prior decision and remove the content.

In their appeal to the Board, the reporting user (who appealed Meta’s decision to leave up the content) stated it was shocking for the company to keep up the content because it was “blatantly using neonazi holocaust denial arguments.” Noting that “millions of Jews, Roma, Disabled, and LGBTQ people were murdered by the Nazi regime,” the reporting user emphasized that this content is hate speech and illegal in Germany.

The user who posted the content claimed in their submission to the Board that they were an “LGBT comedian” who was on a mission to parody the talking points and beliefs of the “alt-right [alternative right].” They said they believed the post was removed for making fun of the “alt right’s beliefs in Holocaust denial” and that their mission was to “uplift marginalized communities.”

6. Meta Submissions

After the Board selected this case, Meta reviewed its original decision and ultimately decided to remove the content for violating its Hate Speech policy. It did not apply a standard strike to the content creator’s account as the content had been posted more than 90 days previously. This is in accordance with Meta’s [strike policy](#). Meta explained to the Board that the specific prohibition of Holocaust denial in the Hate Speech policy was added approximately one month after the user posted the content in question. Meta explained that the second human review of the content, on May 25, 2023, erroneously found the content non-violating as it happened after the policy change. In response to questions from the Board, Meta confirmed that prior to the change, Holocaust denial content would not have been removed, but if it had been coupled with additional hate speech or another violation of the Community Standards, it would have been removed. Meta said the content in this case did not contain any additional hate speech or violations.



Meta noted in its submission to the Board that the content violated the current Hate Speech policy by “denying the existence of the Holocaust.” First, it questions the number of victims, suggesting it is not possible that six million Jewish people were murdered based on supposed population numbers. It also calls into question the existence of crematoria at Auschwitz.

The Board asked Meta 13 questions. These related to the company’s COVID-19 automation policies that led to reports being auto-closed; the policy development process that led to Holocaust denial being prohibited; its enforcement practices related to Holocaust denial content; and the measures that Meta is taking to provide reliable information about the Holocaust and the harms of antisemitism. All questions were answered.

7. Public Comments

The Oversight Board received 35 public comments relevant to this case. Seven comments were submitted from Asia Pacific and Oceania; three from Central and South Asia; four from Europe; one from Latin America and the Caribbean; five from Middle East and North Africa; and 15 from the United States and Canada.

The submissions covered the following themes: the online and offline harms resulting from antisemitic hate speech; social media platforms’ Holocaust denial policies and their enforcement; and how international human rights standards on limiting expression should be applied to moderation of Holocaust denial content. To read public comments submitted for this case, please click [here](#).

8. Oversight Board Analysis

The Board examined whether this content should be removed by analyzing Meta’s content policies, human-rights responsibilities and values. The Board also assessed the implications of this case for Meta’s broader approach to content governance.

The Board selected this case because it provided an opportunity to examine the structural issues that could contribute to this type of content evading detection and removal, and the issue of content removal due to changes in Meta’s policy. It also enabled the Board to evaluate the merits of the Holocaust denial policy in general, under applicable human-rights standards.



8.1 Compliance With Meta’s Content Policies

I. Content Rules

The Board finds that the content in this post violates Meta’s Hate Speech Community Standard, which prohibits Holocaust denial on Facebook and Instagram.

The Board reached out to external experts to clarify how the forms of denial and distortion in this case content fit into racist and antisemitic narratives on Meta’s platforms and more broadly. Experts confirmed that all of the claims in the post were forms of Holocaust denial or distortion: while some of the claims were blatantly untrue, others misrepresented historical facts. Experts also noted that the claims in the content are common antisemitic Holocaust denial tropes on social media. Finally, and as the Brandeis Center noted in their public comment, “[t]he Holocaust was proven beyond a reasonable doubt in front of a duly constituted international court. In its judgment in the case against Major War Criminals of the Nazi regime, the Nuremberg Tribunal considered that the Holocaust had been ‘proved in the greatest detail’” (PC-15024, Louis D. Brandeis Center for Human Rights Under Law).

The Board also commissioned an assessment of Holocaust denial content on Meta’s platforms to understand its prevalence and nature, and the assessment revealed the use of the Squidward meme format to spread various types of antisemitic narratives. The assessment primarily used CrowdTangle, a social media research tool, and was limited to publicly available content. Nonetheless, it provided helpful insight into potential user exposure to Holocaust denial content and confirmed that the content in this case fits into dominant Holocaust denial narratives.

In its Hate Speech Community Standard, Meta explains that it may allow content that would otherwise be prohibited for purposes of “condemnation” or “raising awareness,” or if it is “used self-referentially” or in an “empowering way.” Meta explains that to benefit from these exceptions, it requires people to “clearly indicate their intent.” The Board finds that none of these exceptions applied to this case content.

Additionally, under a heading requiring “additional information and/or context to enforce,” there is also an exception for satire, introduced as the result of a Board recommendation in



the [Two Buttons Meme](#) case. This exception only applies “if the violating elements of the content are being satirized or attributed to something or someone else in order to mock or criticize them.” The content creator in this case claims their post was intended to “parody talking points of the alt-right,” and “uplift marginalized communities.” However, the Board finds no evidence of this stated intent in the post itself. There is none of the exaggeration characteristic of satire in the meme, which replicates typical claims made by Holocaust deniers. Similarly, the cartoon meme style in which the claims are presented are the same as typical Holocaust denial content deployed to attack Jewish people. The assessment the Board commissioned noted that “children’s television cartoon characters are often co-opted, particularly in meme formats, in order to bypass content moderation systems and target younger audiences.” As noted above, Squidward is a children’s cartoon character that is used in multiple antisemitic meme formats. Moreover, the hashtags used do not denote satirical intent, but rather appear to be a further attempt to increase the reach of the content. Finally, the content creator’s comment on their own post, in response to criticism from other users, that the content is “real history” indicates that others did not understand the post to be satirical and shows the user doubling-down on the false claims.

The first human review of this content occurred on October 7, 2020, while the September 23, 2020, Hate Speech policy was still in place and prior to the explicit Holocaust denial prohibition being introduced. The content was also later reviewed by a human reviewer after the prohibition had been introduced, on May 25, 2023. Given that none of the exceptions applied, Meta should have found in the second review that the content violated the current policy on Holocaust denial. As Meta now accepts, the content disputed the number of victims of the Holocaust and the existence of crematoria at Auschwitz. The Board additionally finds that the content calls into question the fact that the Holocaust happened by claiming world leaders’ memoirs didn’t mention it, and that this claim also violates the prohibition on Holocaust denial.

Under the Hate Speech policy prior to the October 2020 changes, the content should have still been removed, as it also violated the pre-existing prohibition on “mocking the concept, events or victims of hate crimes.” To deny and distort key facts of the Holocaust using a cartoon character in the style of a meme was inherently mocking, as it ridicules the Holocaust as a “hate crime,” as well as mocks the memory of its victims.

II. Enforcement Action



The assessment commissioned by the Board reviewed Holocaust denial content on Meta’s platforms and found that determined users try to evade enforcement in various ways, such as by replacing vowels in words with symbols or creating implicit narratives about Holocaust denial that use memes, cartoons and other tropes to relay the same sentiment without directly saying, for example, “the Holocaust didn’t happen.” It also found that “while searches for neutral or factual terms... did yield credible results, other searches for more charged terms led to Holocaust denial content.” The report confirmed the prevalence of claims minimizing the number of Jewish people who were murdered in the Holocaust. Finally, the report noted that Holocaust denial-related content is easier to find and gets more interaction on Instagram than on Facebook.

The assessment also shows a marked decline in content using terms like “Holohoax” and the name of a neo-Nazi propaganda film, “Europa, the Last Battle,” since October 2020, but that there are still gaps in Meta’s removal of Holocaust denial content. As noted by the ADL in its public comment to the Board, “Holocaust denial and distortion continues to be broadcast in mainstream spaces, both on and offline. Despite clear policies that prohibit Holocaust denial and distortion, this antisemitic conspiracy theory still percolates across social media” (PC- 15004, Anti-Defamation League).

The Board notes with concern that the content in this case evaded removal even after Meta changed its policies to explicitly prohibit Holocaust denial, despite two reports being made after the policy change and one being reviewed by a human moderator. As explained below, COVID-19 automation policies led to the automatic closure of one of the reports made on this content after the policy change. Furthermore, as Meta does not require its at-scale reviewers to document the reasons for finding content non-violating, there is no further information about why the human reviewer who reviewed the May 25, 2023 report incorrectly kept the content on the platform. The Board emphasizes that when Meta changes its policies, it is responsible for ensuring that human and automated enforcement of those policies is properly and promptly updated. If content posted prior to policy changes is reported by another user or detected by automation after a policy change that impacts that content, as happened in this case, it should be actioned in accordance with the new policy. That requires updating training materials for human reviewers, as well as classifiers or any other automated tool used to review content on Meta’s platforms, and



ensuring systems are in place to measure the effectiveness of these interventions in operationalizing updates to the Community Standards.

When the Board asked Meta how effective its moderation systems are at removing Holocaust denial content, Meta was not able to provide the Board with data. The Board takes note of Meta’s claimed capacity limitations in measuring both the amount of violating Holocaust denial content on its platforms, and the accuracy of its enforcement, but also understands that these challenges are technically surmountable, if resource intensive. Currently, human reviewers are not given the opportunity to label enforcement data with any granularity. For example, violating content is labelled as “hate speech” rather than as “Holocaust denial.”

The Board recommends that Meta build systems to label enforcement data, including false positives (mistaken removal of non-violating posts) of Holocaust denial content, at a more granular level – especially in view of the real-world consequences of Holocaust denial identified by Meta when it made its policy change. This would enable Meta to measure and report on enforcement accuracy, increasing transparency and potentially improving accuracy. With the limits of human and automated moderation, and the increasing reliance on artificial intelligence to aid content moderation, the Board is interested in how the development of such systems can be shaped to prioritize improving more accurate enforcement at a more granular policy level. In response to the Board's recommendation no. 5 in the Mention of the Taliban in News Reporting case, Meta said that it would develop new tools that would allow it to "gather more granular details about our enforcement of the [Dangerous Organizations and Individuals] news reporting policy allowance." In the Board’s view, this should also be extended to enforcement of the Hate Speech policy.

The Board is also concerned about the application of Meta’s COVID-19 automation policies that were still in force as of May 2023. These led to the automatic closure of one of the reports made on this content after the Hate Speech policy was changed, as well as the automatic closure of the appeal that led to the Board taking on this case. Meta [first announced](#) that it would be sending content reviewers home due to the COVID-19 pandemic in March 2020. In response to questions from the Board, Meta explained that the “policy was created at the beginning of the COVID-19 pandemic in 2020 due to a temporary reduction in human reviewer capacity. This automation policy auto-closed review jobs



based on a variety of conditions and criteria to reduce the volume of reports for human reviewers but kept open [for review] reports that are potentially high risk.”

The user’s appeal was auto-closed in May 2023, more than three years after the COVID-19 pandemic began, and shortly after both the WHO and the United States [declared](#) COVID-19 was no longer a “public health emergency of international concern.” The Board is concerned that measures Meta introduced to handle the pandemic at its outset, which significantly reduced the availability of access to appeal and careful human review, became a new and permanent modus operandi enduring long after circumstances reasonably justified it. During the COVID-19 pandemic, [antisemitism increased](#) and [conspiracy theories circulated](#) claiming that Jewish people were purposefully spreading the virus. There was a pressing need for Meta to prioritize the review and removal of hate speech, given the severe impacts of such speech on individuals’ rights, as soon as the circumstances of this emergency allowed. The Board is concerned that what was introduced as a pandemic contingency could be extended for a significant period of time when the necessity of such a significant scaling back of the kind of careful human review necessary to implement Meta’s detailed and sensitive policies is not demonstrated, and recommends that Meta restore review of content moderation decisions as soon as possible, and publish information in its Transparency Center when it does so.

8.2 Compliance With Meta’s Human-Rights Responsibilities

Freedom of Expression (Article 19 ICCPR)

Article 19 of the ICCPR provides for broad protection of the right to freedom of expression, including discussions on matters of history. The Human Rights Committee has said that the scope of this right “embraces even expression that may be regarded as deeply offensive, although such expression may be restricted in accordance with the provisions of article 19, paragraph 3 and article 20,” ([General Comment No. 34](#), para. 11). 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, both in relation to the individual content decision under review and what this says about Meta’s broader approach to content governance.



Additionally, the ICCPR requires states to prohibit advocacy of racial hatred that constitutes incitement to hostility, discrimination or violence (Article 20, para. 2, ICCPR). As the UN Special Rapporteur on freedom of 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](#)). Meta has the responsibility to prevent and mitigate incitement on its platforms.

Public comments in this case reflect diverging views on how international human-rights standards on limiting expression should be applied to the moderation of Holocaust denial content online. Several public comments argued that Meta’s human-rights responsibilities require such content to be removed (see PC-15023, American Jewish Committee and its Jacob Blaustein Institute for the Advancement of Human Rights; PC-15024, Louis D. Brandeis Center for Human Rights Under Law; and PC-15018, Prof. Yuval Shany of the Hebrew University of Jerusalem Faculty of Law). Others argued that Meta should address the lack of specificity in the policy by defining Holocaust denial, making clearer the prohibition aims at addressing antisemitism, as well as improve training of human reviewers (see PC-15034, University of California, Irvine – International Justice Clinic). Finally, some public comments argued that Holocaust denial content should only be removed when it constitutes direct incitement to violence under Article 20, para. 2 of the ICCPR (see PC-15022, Future of Free Speech Project).

I. Legality (Clarity and Accessibility of the Rules)

The Board finds that the current Hate Speech policy prohibition on Holocaust denial is sufficiently clear to satisfy the legality standard. Since its revision in October 2020, the Hate Speech Community Standard clearly states that content denying the Holocaust is not allowed. However, the Board notes that the language is less clear than when it was first introduced, in two ways. First, as noted above, UN resolution 76/250 specifically urges social media companies to address Holocaust denial *or* distortion [emphasis added]. That means the removal of the word “distortion” (originally included alongside “denial”) in 2022 lessened the policy’s conformity with UN recommendations. Second, the placement of the policy line under the prohibition on “Harmful stereotypes historically linked to intimidation, exclusion, or violence on the basis of a protected characteristic,” reduces the policy’s clarity. Holocaust denial is linked to antisemitic stereotypes, but not all instances will necessarily be an example of direct stereotyping.



The Hate Speech policy prior to the October 2020 revisions did not expressly prohibit Holocaust denial, but the prohibition on “mocking the concept, events or victims of hate crimes” did, in the Board’s view, cover most instances of Holocaust denial, even if it did not address fully the nature of the Holocaust.

Notwithstanding that the current policy on Holocaust denial is expressly included in the Facebook Community Standards, the same is not true of the Instagram Community Guidelines, in which Holocaust denial is not mentioned at all. The Board emphasizes that it has asked Meta in several recommendations to align its Instagram and Facebook standards and distinguish where there are inconsistencies. Meta has committed to implementing these recommendations fully, but it also explains in its Transparency Center that it does “not believe adding a short explanation to the Community Guidelines introduction will fully address the board’s recommendation and may lead to further confusion. Instead, we are working to update the Instagram Community Guidelines so that they are consistent with the Facebook Community Standards in all of the shared policy areas.” In its quarterly update, Meta said this is a key priority but has had to be deprioritized because of regulatory compliance work. Meta will not complete this recommendation this year and expects to have an update on the progress in Q2 2024. Noting that its commissioned research and civil society investigations indicate that Holocaust denial is more prevalent on Instagram, the Board reiterates its prior recommendation and urges Meta to continue to communicate any delays and implement any short-term policy solutions available to bring more clarity to Instagram users, in particular on the issue of Holocaust denial.

Content is accessible on Meta’s platforms on a continuing basis and content moderation policies are applied on a continuing basis. Therefore, Meta removing old posts still hosted on Facebook or Instagram, after a rule change that clearly prohibits that content, does not violate the requirements of legality. Rather, continuous publication of the content that Meta hosts after a substantive policy change or clarification when it comes to Tier 1 (and in other situations where human life is at risk) necessitates removal, even for posts that pre-date the introduction of new rules. Meta does not count strikes “on violating content posted over 90 days ago for most violations or over 4 years ago for more severe violations.” This means that in most cases, there would also be no penalty for previously permitted content that later comes to violate new rules. However, the [strikes policy](#) means that users could incur penalties where Meta changes a rule and subsequently enforces it against



content posted up to 90 days prior to the rule change. The Board emphasizes that while it is consistent with the principle of legality to remove content after a rule change in the specific context of social media, for the reasons outlined above, it is not appropriate to apply retroactive punishment in the form of strikes when removing content that was permitted when it was posted.

II. Legitimate Aim

In numerous cases, the Board has recognized that Meta’s Hate Speech Community Standard pursues the legitimate aim of protecting the rights of others. Meta explicitly states that it does not allow hate speech because it “creates an environment of intimidation and exclusion, and in some cases may promote offline violence.” Meta also noted similar aims when it announced the introduction of a specific prohibition on Holocaust denial. Numerous public comments noted this addition was necessary to safeguard against increased incitement to violence and hostility (PC-15023, American Jewish Committee and its Jacob Blaustein Institute for the Advancement of Human Rights), emphasizing that Holocaust denial and distortion amounts to a discriminatory attack against Jewish people and promotes antisemitic stereotypes, often connected to and spread during antisemitic hate crimes.

It is important to understand Holocaust denial as a constitutive element of antisemitism that is discriminatory in its consequences. The denial of the Holocaust amounts to the denial of “barbarous acts which have outraged the conscience of mankind,” as described by the Universal Declaration of Human Rights (see also UN General Assembly Resolution [76/250](#)). The Hate Speech Community Standard and its prohibition on Holocaust denial pursues the legitimate aim of respecting the rights to equality and non-discrimination of Jewish people as well as their right to freedom of expression. Allowing such hate speech would create an environment of intimidation that effectively excludes Jewish people from Meta’s platforms [See, e. g, PC-15021, Monika Hübscher, noting that “individuals impacted by antisemitic hate speech on social media describe the attacks in a language that equals the depictions of physical acts. Exposure to hate on social networks can lead to feelings of fear, insecurity, heightened anxiety, and even sleep disturbance”]. Meta’s prohibition on Holocaust denial also serves the legitimate aim of respecting the right to reputation and the dignity and memory of those who perished in the most inhumane circumstances and the rights of their relatives. Such hate speech is a fundamental attack on the dignity of human beings (see also Universal Declaration of Human Rights, Article 1).



III. Necessity and Proportionality

Meta's decision to ban Holocaust denial is consistent with its human-rights responsibilities. The Board notes that Meta's responsibilities to remove hate speech in the form of Holocaust denial can be considered necessary and proportionate in numerous ways. Under ICCPR Article 19, para. 3, necessity requires that restrictions on expression "must be appropriate to achieve their protective function." The removal of the content would not be necessary "if the protection could be achieved in other ways that do not restrict freedom of expression" ([General Comment No. 34](#), para. 33). Proportionality requires that any restriction "must be the least intrusive instrument amongst those which might achieve their protective function" ([General Comment No. 34](#), para. 34).

The Board considers that there are different ways to approach content that denies the Holocaust. While the majority of Board members consider - for various reasons explained below - that the prohibition on Holocaust denial satisfies the principle of necessity and proportionality, a minority considers that Meta did not meet the conditions for establishing this prohibition.

For the majority, UN General Comment No. 34 does not invalidate prohibitions on Holocaust denial that are specific to the regulation of hate speech, as Meta's prohibition is, when such denial is understood as an attack against a protected group. Meta's rule expressly prohibiting Holocaust denial as hate speech was a response to an alarming rise in the dissemination of such antisemitic content online that was internationally denounced; the on and offline harm that such hate speech causes; and the staggering ignorance about the commission of these heinous crimes of the Holocaust that offend the conscience of humanity and whose veracity has been conclusively demonstrated. The Board notes that the prohibition is also responsive to UN General Assembly Resolution 76/250, which "urges ...social media companies to take active measures to combat antisemitism and Holocaust denial or distortion by means of information and communications technologies and to facilitate reporting of such content." The majority notes that Meta's prohibition is also not absolute, as specific exceptions exist to allow condemnation, awareness raising and satire, as well as broader exceptions such as the newsworthiness allowance. The ban on Holocaust denial is therefore in conformity with the ICCPR and the obligations expressed in Article 4 (a) of the International Convention on the Elimination of All Forms of Racial Discrimination. Holocaust denial is "a dissemination of



ideas based on racial hatred,” given that “Holocaust denial in its various forms is an expression of antisemitism” (see also UN General Assembly Resolution [76/250](#)). Furthermore, in the above-mentioned context, the post, by denying the facts of the Holocaust, may contribute to the creation of an extremely hostile environment on the platform, causing exclusion of the impacted communities, and profound pain and suffering. Therefore, there is “a direct and immediate connection between the expression and the threat” to voice, dignity, safety and reputation of others that justifies the prohibition in the sense required by General Comment 34, at para. 35.

For some members of the majority, there are additional reasons to support Meta’s prohibition. A legally proven fact cannot be the subject matter of divergent opinions when such lies have directly harmful consequences on others’ rights to be protected from violence and discrimination. The presentation of Holocaust denial as opinion about “historical facts” is therefore an abuse of the right to freedom of expression. These same members note that in *Faurisson v. France*, (550/1993) the UN Human Rights Committee found that a ban on Holocaust denial complied with the requirements of Article 19, para. 3. The Committee came to this conclusion in the context of the French Gaysot Act, which made it illegal to question the existence or size of the crimes against humanity recognized in the Charter of the Nuremberg Tribunal. The Committee’s decision, which relates to enforcement of a law that would seemingly prohibit the content under consideration in this case, supports the Board’s conclusion that Meta’s eventual removal of the post was permissible under international human rights law.

For other members of the majority, who depart from considering the *Faurisson* case to be a currently valid doctrine, the company's decision is consistent with the principles of necessity and proportionality for different reasons summarized below and arising from the Board's precedents. In previous cases, the Board has agreed with the UN Special Rapporteur on freedom of expression that although some restrictions (such as general bans on certain speech) would generally not be consistent with governmental human rights obligations (particularly if enforced through criminal or civil penalties), Meta may prohibit such speech provided that it demonstrates the necessity and proportionality of the restriction (see [South Africa Slurs](#) decision and [Zwarte Piet](#) decision). In these cases, companies should "give a reasoned explanation of the policy difference in advance, in a way that articulates the variation" ([A/74/486](#), para. 48; [A/HRC/38/35](#), para. 28). For a prohibition of this kind to be compatible with Meta's responsibilities, it is necessary that it



be based on a human rights analysis demonstrating that the policy is in pursuit of a legitimate aim; is useful, necessary and proportionate to achieve that aim (see [South Africa Slurs](#) decision); and that the prohibition should be periodically reviewed to ensure that the need persists (see [Removal of COVID-19 Misinformation policy advisory opinion](#)).

For these members of the majority, these conditions are met as demonstrated by the evidence the Board found and summarized in earlier sections of this decision, particularly, in the alarming rise of antisemitism globally and the growth online and offline of antisemitic violence.

The Board agrees that there are different forms of intervention that social media platforms such as Meta can deploy besides content removal to address hate speech against Jewish people. The UN Special Rapporteur on freedom of opinion and expression has recommended that social media companies should consider a range of possible responses to problematic content beyond removal to ensure restrictions are narrowly tailored, including geo-blocking, reducing amplification, warning labels and promoting counter-speech ([A/74/486](#), para. 51). The Board welcomes [various initiatives from Meta](#) to counter antisemitism, in addition to removal of violating content, including educating people about the Holocaust, directing people to credible information off Facebook if they search for terms associated with the Holocaust or its denial on its platforms, and engaging with organizations and institutions that work on combating hate and antisemitism. The Board encourages Meta to roll these initiatives out uniformly across Instagram and explore targeting them at people who violate the Holocaust denial policy.

For the majority, given the evidence of the negative impact of Holocaust denial on users of Meta's platforms, these measures, while valuable, cannot fully protect Jewish people from discrimination and violence. As public comments also note, "less severe interventions than removal of Holocaust denial content, such as labels, warning screens, or other measures to reduce dissemination, may be useful but would not provide the same protection [as removal]" (PC-15023, American Jewish Committee and its Jacob Blaustein Institute for the Advancement of Human Rights). In the absence of less intrusive means to effectively combat hate speech against Jewish people, the majority finds the Holocaust denial prohibition meets the requirements of necessity and proportionality.



While a minority of Board Members also firmly condemns Holocaust denial and believes it should be addressed by social media companies, they find the majority’s necessity and proportionality analysis is out of step with the UN human rights mechanisms’ approach to freedom of expression over the last 10 years. First, with regard to reliance on the Human Rights Committee’s 1986 Faurisson case in justifying the removal as necessary and proportionate, the minority highlighted (as did PC-15022, Future of Free Speech Project) that the lead drafter of General Comment 34 confirmed that Faurisson was effectively overruled by General Comment 34, which was adopted in 2011 ([Michael O’Flaherty, Freedom of Expression: Article 19 of the ICCPR and the Human Rights Committee’s General Comment 34](#), 12 Hum. Rts. L.Rev. 627, 653 (2012)). Paragraph 49 of General Comment 34 states that the ICCPR does not permit the general prohibition of expressions of erroneous opinions about historical facts. Any restrictions on expression must meet the strict tests of necessity and proportionality, which require considering likely and imminent harm. The minority finds that the reliance on Article 4 of the ICERD is misplaced as the Committee on the Elimination of Racial Discrimination (CERD, which is charged with monitoring implementation of the ICERD) specifically addressed the topic of genocide denial, stating it should only be banned when the statements “clearly constitute *incitement* to racial violence or hatred. The Committee also underlines that ‘the expression of opinions about historical facts’ should not be prohibited or punished,” [CERD General Recommendation No. 35](#), para 14) [emphasis added].

This minority of Board Members is not convinced that content removal is the least intrusive means available to Meta to address antisemitism, and that Meta’s failure to demonstrate otherwise does not satisfy the requirement of necessity and proportionality. The Special Rapporteur has stated “just as States should evaluate whether a limitation on speech is the least restrictive approach, so too should companies carry out this kind of evaluation. And, in carrying out the evaluation, companies should *bear the burden of publicly demonstrating necessity and proportionality*” ([A/74/486](#), para. 51) [emphasis added]. For this minority, Meta should have publicly demonstrated why removal of such posts is the least intrusive means of the many tools it has at its disposal to avert likely near-term harms, such as discrimination or violence. If it cannot provide such a justification, then Meta should be transparent in acknowledging that its speech rules depart from UN human-rights standards and provide a public justification for doing so. This minority believes that the Board would then be positioned to consider Meta’s public justification and a public dialogue would ensue without distorting existing UN human-rights standards.



9. Oversight Board Decision

The Oversight Board overturns Meta’s original decision to leave up the content.

10. Recommendations

Enforcement

1. To ensure that the Holocaust denial policy is accurately enforced, Meta should take the technical steps to ensure that it is sufficiently and systematically measuring the accuracy of its enforcement of Holocaust denial content. This includes gathering more granular details about its enforcement of this content, as Meta has done in implementing the Mention of the Taliban in News Reporting recommendation no. 5.

The Board will consider this recommendation implemented when Meta provides the Board with its first analysis of enforcement accuracy of Holocaust denial content.

Transparency

2. To provide greater transparency that Meta’s appeals capacity is restored to pre-pandemic levels, Meta should publicly confirm whether it has fully ended all COVID-19 automation policies put in place during the COVID-19 pandemic.

The Board will consider this recommendation implemented when Meta publishes information publicly on each COVID-19 automation policy and when each was ended or will end.

The Oversight Board also reiterates the importance of its previous recommendations calling for alignment of the Instagram Community Guidelines and Facebook Community Standards, noting the relevance of these recommendations to the issue of Holocaust denial (recommendation no. 7 and 9 from the Breast Cancer Symptoms and Nudity case; recommendation no. 10 from the Öcalan’s Isolation case; no. 1 from the Ayahuasca Brew case; and recommendation no. 9 from the Sharing Private Residential Information policy advisory opinion). In line with those recommendations, Meta should continue to



communicate delays in aligning these rules, and it should implement any short-term solutions to bring clarity to Instagram users.

***Procedural Note:**

The Oversight Board's decisions are prepared by panels of five Members and approved by a majority of the Board. Board decisions do not necessarily represent the personal views of all Members.

For this case decision, independent research was commissioned on behalf of the Board. The Board was assisted by an independent research institute headquartered at the University of Gothenburg, which draws on a team of over 50 social scientists on six continents, as well as more than 3,200 country experts from around the world. The Board was also assisted by Duco Advisors, an advisory firm focusing on the intersection of geopolitics, trust and safety, and technology. Memetica, an organization that engages in open-source research on social media trends, also provided analysis.