Politician's Comments on Demographic Changes

2023-033-FB-UA

Summary

The Oversight Board has upheld Meta's decision to leave up a video clip in which French politician Éric Zemmour discusses demographic changes in Europe and Africa. The content does not violate the Hate Speech Community Standard since there is no direct attack on people based on a protected characteristic such as race, ethnicity or national origin. The majority of the Board find that leaving up the content is consistent with Meta's human rights responsibilities. However, the Board recommends that Meta should publicly clarify how it distinguishes immigration-related discussions from harmful speech, including hateful conspiracy theories, targeting people based on their migratory status.

About the Case

In July 2023, a video clip in which French politician Éric Zemmour discusses demographic changes in Europe and Africa was posted on his official Facebook page by a user who is the page's administrator. The clip is part of a longer video interview with the politician. In the video, Zemmour states: "Since the start of the 20th century, there has been a population explosion in Africa." He goes on to say that while the European population has stayed roughly the same at around 400 million people, the African population has increased to 1.5 billion people, "so the power balance has shifted." The post's caption, in French, says that in the 1900s, "when there were four Europeans for one African, [Europe] colonized Africa," and now "there are four Africans for one European and Africa colonizes Europe." Zemmour's Facebook page has about 300,000 followers while this post had been viewed about 40,000 times as of January 2024.

Zemmour has been the subject of multiple legal proceedings, with more than one conviction in France for inciting racial hatred and making racially insulting comments about Muslims, Africans and Black people. He ran for president in 2022 but did not progress beyond the first round. Central to his electoral campaigning is the Great Replacement Theory, which argues that white European populations are being deliberately replaced ethnically and culturally through migration and the growth of minority communities. Linguistic experts note the

theory and terms associated with it "incite racism, hatred and violence targeting the immigrants, non-white Europeans and target Muslims specifically." The video in the post does not specifically mention the theory.

Two users reported the content for violating Meta's Hate Speech policy but since the reports were not prioritized for review in a 48-hour period, they were both automatically closed. Reports are prioritized by Meta's automated systems according to the severity of the predicted violation, the content's virality (number of views) and likelihood of a violation. One of the users then appealed to Meta, which led to one of the company's human reviewers deciding the content did not violate Meta's rules. The user then appealed to the Board.

Key Findings

The majority of the Board conclude the content does not violate Meta's Hate Speech Community Standard. The video clip contains an example of protected (albeit controversial) expression of opinion on immigration and does not contain any call for violence, nor does it direct dehumanizing or hateful language towards vulnerable groups. While Zemmour has been prosecuted for use of hateful language in the past, and themes in this video are similar to the Great Replacement Theory, these facts do not justify removal of a post that does not violate Meta's standards.

For there to have been a violation, the post would have had to include a "direct attack," specifically calling for the "exclusion or segregation" of a "protected characteristic" group. Since Zemmour's comments do not contain any direct attack, and there is neither an explicit call to exclude any group from Europe nor any statement about Africans tantamount to a harmful stereotype, slur or any other direct attack, they do not break Meta's Hate Speech rules. The policy rationale also makes it clear that Meta allows "commentary on and criticism of immigration policies," although what is not shared publicly is that the company allows calls for exclusion when immigration policies are being discussed.

However, the Board does find it concerning that Meta does not consider Africans a protected characteristic group, given the fact that national origin, race and religion are protected both under Meta's policies and international human rights law. Africans are mentioned throughout the content and, in this video, serve as a proxy for non-white Africans.

The Board also considered the relevance of the Dangerous Organizations and Individuals policy to this case. However, the majority find the post does not violate this policy because there are not enough elements to review it as part of a wider Violence-Inducing Conspiracy Network. Meta defines these networks as non-state actors who share the same mission statement, promote unfounded theories claiming that secret plots by powerful actors are behind social and political problems, and who are directly linked to a pattern of offline harm.

A minority of Board Members find that Meta's approach to content spreading harmful conspiracy theories is inconsistent with the aims of the policies it has designed to prevent an environment of exclusion affecting protected minorities, both online and offline. Under these policies, content involving certain other conspiracy narratives is moderated to protect threatened minority groups. While these Board Members believe that criticism of issues like immigration should be allowed, it is precisely because evidence-based discussion on this topic is so relevant that the spread of such conspiracy theories, such as the Great Replacement Theory, can be harmful. It is not individual pieces of content but the combined effects of such content shared on a large scale and at high speeds that causes the greatest challenge to social media companies. Therefore, Meta needs to reformulate its policies so that its services are not misused by those who promote conspiracy theories causing online and offline harm.

Meta has undertaken research into a policy line that could address hateful conspiracy theories but the company decided this would ultimately lead to removal of too much political speech. The Board is concerned about the lack of information that Meta shared on this process.

The Oversight Board's Decision

The Oversight Board has upheld Meta's decision to leave up the post.

The Board recommends that Meta:

- Provide greater detail in the language of its Hate Speech Community Standard on how it distinguishes immigration-related discussions from harmful speech targeting people based on their migratory status. This includes explaining how the company handles content spreading hateful conspiracy theories, so that users can understand

how Meta protects political speech on immigration while addressing the potential offline harms of these theories.

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

Full Case Decision

1. Decision Summary

The Oversight Board upholds Meta's decision to leave up a post on French politician Éric Zemmour's official Facebook page that contains a video of Mr. Zemmour being interviewed, in which he discusses demographic changes in Europe and Africa. The Board finds the content does not violate Meta's Hate Speech Community Standard because it does not directly attack people on the basis of protected characteristics, including race, ethnicity and national origin. The majority of the Board find that Meta's decision to keep the content on Facebook is consistent with its human rights responsibilities. A minority of Board Members find that Meta's policies are inadequate to meet Meta's human rights responsibilities to address the significant threat posed by harmful and exclusionary conspiracy theories such as the Great Replacement Theory. The Board recommends that Meta publicly clarify how it handles content spreading hateful conspiracy theories given the need to protect speech about immigration while addressing the potential offline harms of such harmful conspiracy theories.

2. Case Description and Background

On July 7, 2023, a user posted a video on the official, verified Facebook page of French politician Éric Zemmour. In the video, which is a 50-second clip of a longer interview conducted in French, Zemmour discusses demographic changes in Europe and Africa. The user who posted the video was an administrator of the page, which has about 300,000 followers. Zemmour was a candidate in the 2022 French presidential election and won around 7% of the votes in the first round, but did not advance any further. Before running for office, Zemmour was a regular columnist at Le Figaro and other newspapers, as well as an outspoken TV commentator famed for his provocations on Islam, immigration and women. As explained in greater detail below, he has been involved in multiple legal proceedings and convicted in some

of them on account of these comments. Although Meta did not consider the user who posted the video to be a public figure, the company did consider Zemmour a public figure.

In the video, Zemmour states that: "Since the start of the 20th century, there has been a population explosion in Africa." He states that while the European population has stayed roughly the same at around 400 million people, the African population has increased to 1.5 billion people, "so the power balance has shifted." The caption in French repeats the claims in the video, stating that in the 1900s, "when there were four Europeans for one African, [Europe] colonized Africa," and now "there are four Africans for one European and Africa colonizes Europe." These figures are compared to figures available from United Nations bodies provided below. Additionally, the Board's majority position on these numbers is described in greater detail in Section 8.2 below.

When this case was announced by the Board on November 28, 2023, the content had been viewed around 20,000 times. As of January 2024, the content had been viewed about 40,000 times and had fewer than 1,000 reactions, the majority of which were "likes," followed by "love" and "Haha."

On July 9, 2023, two users separately reported the content as violating Meta's Hate Speech policy. The company automatically closed both reports because they were not prioritized for review in a 48-hour period. Meta explained that reports are dynamically prioritized for review based on factors such as the severity of the predicted violation, the content's virality (number of views the content has had) and the likelihood that the content will violate the company's policies. The content was not removed and stayed on the platform.

On July 11, 2023, the first person who reported the content appealed Meta's decision. The appeal was assessed by a human reviewer who upheld Meta's original decision to keep the content up. The reporting user then appealed to the Oversight Board.

Ten days before the content was posted, the fatal shooting of Nahel Merzouk, a French 17-year-old of Moroccan and Algerian descent, who died after two police officers shot him at point-blank range in a suburb of Paris on June 27, 2023, sparked widespread riots and violent protests in France. The protests, which were ongoing when the content was posted, were directed at police violence and the perceived systemic racial discrimination of policing in

France. These protests were the most recent [in a long series](#) of protests about police violence that, it is claimed, often targets immigrants of African origin and other marginalized communities in France.

According to the [European Commission against Racism and Intolerance](#) (ECRI), the main victims of racism in France are immigrants, especially those of African origin and their descendants. In 2022, the Committee on the Elimination of Racial Discrimination (CERD) [urged France](#) to redouble its efforts to effectively prevent and combat racist hate speech and said that "despite the State party's efforts… the Committee remains concerned at how persistent and widespread racist and discriminatory discourse is, especially in the media and on the Internet." It is also "concerned at some political leaders' racist remarks with regard to certain ethnic minorities, in particular Roma, Travellers, Africans, persons of African descent, persons of Arab origin and non-citizens," (CERD/C/FRA/CO/22-23, para. 11).

According to a [1999 report](#) from the Department of Social and Economic Affairs of the United Nations, the estimated population of Africa in the year 1900 was 133 million. Data from the [United Nations in 2022](#) estimates that the 2021 population of Africa was around 1.4 billion. [It also projects](#) that by 2050, the estimated population of Africa could be close to 2.5 billion. According to the same 1999 report, the population of Europe in 1900 was approximately 408 million. The Department of Economic and Social Affairs of the [United Nations estimates](#) that the population of Europe in 2021 was approximately 745 million and [will decline](#) to approximately 716 million by 2050.

In France, as in many parts of the world, the arrival of large numbers of migrants from other countries has become one of the most salient topics of political debate. As of September 2023, approximately [700,000 refugees](#) and asylum seekers were located in France, making it the third-largest host country in the European Union. As the Board reviewed this case, France experienced large-scale protests and heated public debate on migration amidst the Parliamentary passage of a [new immigration bill](#) that, among other things, sets migration quotas and tightens the rules around family reunification and access to social benefits. Zemmour and his party [have been very active](#) in these discussions, advocating for immigration restrictions.

Zemmour has been the subject of [multiple legal proceedings](#) and has been convicted several times by French courts for inciting racial hatred and making racially insulting comments in recent years, as a result of his statements about Muslims, Africans, Black people and LGBTQIA+ people. Zemmour [was convicted](#) of [incitement to racial hatred](#) for comments he made in 2011 on television in which he said "most dealers are blacks and Arabs. That's a fact." More recently, a court found [Zemmour guilty](#) of inciting racial hatred in 2020 and fined him 10,000 euros [for stating](#) that child migrants are "thieves, killers, they're rapists. That's all they are. We should send them back." He has exhausted his right to appeal in another case in which he was convicted and fined by the correctional court for inciting discrimination and religious hatred against the French Muslim community. The conviction was based on statements he made on a television show in 2016 that Muslims should be given "the choice between Islam and France" and that "for thirty years we have been experiencing an invasion, a colonization (…) it is also the fight to Islamize a territory which is not, which is normally a non-Islamized land." In December 2022, the [European Court of Human Rights](#) held that the conviction did not violate Zemmour's right to freedom of expression.

Although the content in this case does not explicitly mention the Great Replacement Theory, the concept is central to Zemmour's political ideology [and featured heavily in his presidential campaign,](#) during which he promised[, if elected, to create a "Ministry of Remigration."](#) He also stated that he would "send back a million" foreigners in five years. According to independent research commissioned by the Board, which was conducted by experts in conspiracy theories and French politics, social media trends, and linguistics, proponents of [the Great Replacement Theory](#) argue that white European populations are being deliberately replaced ethnically and culturally through migration and the growth of minority communities. It insists that contemporary migration of non-white (and predominantly Muslim) people from non-European countries (mostly, in Africa and Asia) to Europe is a form of demographic warfare. The Board's experts emphasized that migration and the increase in migration is not factually disputed. Rather, it is the insistence that there is an actual plot or conspiracy to bring non-whites into Europe in order to replace or reduce the proportion of white populations that marks the Great Replacement Theory as conspiratorial. Linguistic experts consulted by the Board explained that the Great Replacement Theory and terms associated with it "incite racism, hatred and violence targeting the immigrants, non-white Europeans and target Muslims specifically." A [report](#) by the European Union's Radicalization Awareness Network notes that the anti-Semitic, anti-Muslim and overall anti-immigration sentiment spread by people advancing the Great Replacement Theory has informed the selection of targets by several high-profile solo

attackers in Europe in recent years. The Board's commissioned research also indicated that the theory has inspired myriad violent incidents around the world in recent years, including the [mass shooting in Christchurch](), New Zealand, in which 51 Muslims were killed.

A minority of the Board also considers the fact that violent far-right protests have been on the rise in France over the past year as important context. Following the [fatal stabbing]() of a teenager during a festive gathering on November 18 in Crépol, a rural community in France, activists and right-wing parties led [violent protests]() in which protestors physically clashed with the police. They alleged that immigrants and minorities were responsible, despite the fact that of the [nine people arrested]() in connection with the stabbing, eight were French, and one Italian. Interior Minister Gérald Darmanin [said]() that militia members "seek to attack Arabs, people with different skin colors, speak of their nostalgia for the Third Reich." French Green Party politician Sandrine Rousseau [compared]() these protests to *ratonnades,* physical violence carried out against an ethnic minority or a social group, predominantly against people of North African origin. The most notable *[ratonnade]()*, often associated with the popularization of the term, occurred on October 17, 1961 during peaceful protests by Algerians in which 200 Algerians were killed during an outburst of police violence. The word has come up repeatedly in different contexts in France since then. For example, in December 2022 French politicians joined social media users in denouncing street violence, [comparing it to]() *ratonnades* after the France-Morocco World Cup match.

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 the Board's 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.

4. Sources of Authority and Guidance

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

I. Oversight Board Decisions

- Holocaust Denial
- Communal Violence in Indian State of Odisha
- Knin Cartoon
- Shared Al Jazeera Post
- Former President Trump's Suspension
- Depiction of Zwarte Piet

II. Meta's Content Policies

Hate Speech

The policy rationale for the Hate Speech Community Standard explains that hate speech, defined as a direct attack against people on the basis of protected characteristics, is not allowed on the platform "because it creates an environment of intimidation and exclusion and, in some cases, may promote real-world violence." The policy lists as protected characteristics, among others, race, ethnicity, national origin and religious affiliation. The policy explains that "attacks are separated into two tiers of severity," with Tier 1 attacks being more severe. The rationale for the Hate Speech Community Standard also explains that the policy "protect(s) refugees, migrants, immigrants and asylum seekers from most severe attacks," but that Meta allows "commentary on and criticism of immigration policies." Meta's internal guidance to content moderators elaborates on this, explaining that it considers migrants, immigrants, refugees and asylum-seekers status as quasi-protected. That means Meta protects them from Tier 1 attacks but not from Tier 2 attacks under the Hate Speech policy.

The policy, which previously had three different attack tiers but now only has two of them, currently prohibits as a Tier 2 attack, among other types of content, "exclusion or segregation in the form of calls for action, statements of intent, aspirational or conditional statements, or statements advocating or supporting defined as: [...] explicit exclusion, which means things

like expelling certain groups or saying they are not allowed." Meta declined the Board's request to publish further information about its internal guidance to content reviewers on this point.

In a 2017 Newsroom post entitled "Hard Questions: Who Should Decide What Is Hate Speech in an Online Global Community?," which is linked at the bottom of the rationale for the Hate Speech Community Standard with the text "Learn more about our approach to hate speech," Meta recognized that policy debates on immigration often become "a debate over hate speech, as two sides adopt inflammatory language." The company said that after reviewing posts on Facebook about the migration debate globally, it "decided to develop new guidelines to remove calls for violence against migrants or dehumanizing references to them — such as comparisons to animals, to filth or to trash." The company left in place, however, "the ability for people to express their views on immigration itself," given it is "deeply committed to making sure Facebook remains a place for legitimate debate."

Dangerous Organizations and Individuals

The rationale for the Dangerous Organizations and Individuals Community Standard states that Meta does not allow organizations or individuals that proclaim a violent mission or are engaged in violence to have a presence on Meta's platforms. It also explains that Meta assesses these entities "based on their behavior both online and offline – most significantly, their ties to violence."

The Dangerous Organizations and Individuals Community Standard explains that Meta prohibits the presence of Violence-Inducing Conspiracy Networks, currently defined as non-state actors that are: (i) "identified by a name, mission statement, symbol or shared lexicon"; (ii) "promote unfounded theories that attempt to explain the ultimate causes of significant social and political problems, events and circumstances with claims of secret plots by two or more powerful actors"; and (iii) "have explicitly advocated for or have been directly linked to a pattern of offline physical harm by adherents motivated by the desire to draw attention to or redress the supposed harms identified in the unfounded theories promoted by the network."

The Board's analysis of the content policies was also informed by Meta's commitment to voice, which the company describes as "paramount" as well as 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 right to freedom of 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).
- Equality and non-discrimination: Article 2, para. 1 and Article 26, ICCPR; Article 2, International Convention on the Elimination of All Forms of Racial Discrimination (ICERD); UN General Assembly, resolution A/RES/73/195.
- Right to life: Article 6, ICCPR.

5. User Submissions

The Board received a submission from the user who reported the content and appealed Meta's decision to keep it up, as part of their appeal to the Board. In the submission, the appealing user says that Zemmour is explaining both colonization and migration in terms of overpopulation only, which the user classified as "fake news."

6. Meta's Submissions

After the Board selected this case, Meta reviewed the post against the Hate Speech policy with subject-matter experts and determined that its original decision to leave up the content was correct. Meta did not provide further information on the specific remit or knowledge areas of the experts conducting this additional review. Meta emphasized that, for a piece of content to be considered as violating, the policy requires both a protected characteristic group and a direct attack – and that the claims about population changes and colonization lacked those elements. Meta explained that it does not consider the allegation that one group is "colonizing" a place to be an attack in and of itself so long as it does not amount to a call for exclusion, and emphasized that it "want[s] to allow citizens to discuss the laws and policies of their nations so

long as this discussion does not constitute attacks against vulnerable groups who may be the subject of those laws." Finally, Meta explained that the content does not identify a protected characteristic group because Zemmour refers to "Africa," a continent and its countries, and that the "Hate Speech policy does not protect countries or institutions from attacks."

Meta refused to lift confidentiality related to the company's policy development process on harmful conspiracy theories. Meta instead stated: "We have considered policy options specific to content discussing conspiracy theories that does not otherwise violate our existing policies. However, we have concluded that, for the time being, implementing any of the options would risk removing a significant amount of political speech."

The Board asked Meta eight questions in writing. Questions covered Meta's policy development in relation to the Great Replacement Theory; the applicability of various Hate Speech and Dangerous Organizations and Individuals policy lines; and the violation history for the Facebook page and posting user. Meta answered six of the Board's questions, with two not answered satisfactorily. After Meta did not provide sufficient detail in response to the Board's initial question about policy development in relation to the Great Replacement Theory, the Board asked a follow-up question to which Meta provided additional but still less than comprehensive information.

7. Public Comments

The Oversight Board received 15 public comments. Seven of the comments were submitted from the United States and Canada, three from Europe, two from Central and South Asia, one from Asia Pacific and Oceania, one from the Middle East and North Africa, and one Sub-Saharan Africa. This total includes public comments that were either duplicates or were submitted with consent to publish but did not meet the Board's conditions for publication. Public comments can be submitted to the Board with or without consent to publish and with or without consent to attribute (i.e., anonymously).

The submissions mainly covered two themes. First, several comments emphasized that removing the content under review in this case would be tantamount to censorship, and could even "serve to increase the anger of the citizens who feel their voices will not be heard," (PC-22009). Second, two organizations submitted comments emphasizing the negative offline

impact of this type of content and, specifically, the Great Replacement Theory. Both of these comments argued that there is a link between the Christchurch massacre and the theory (PC-22013, Digital Rights Foundation; PC-22014, Global Project Against Hate and Extremism).

To read public comments submitted for this case, please click [here](#).

8. Oversight Board Analysis

The Board analyzed Meta's content policies, human rights responsibilities and values to determine whether the content in this case should be removed. The Board also assessed the implications of this case for Meta's broader approach to content governance.

The Board selected this case as an opportunity to review Meta's approach to content targeting migrants in the context of increasingly global anti-immigrant rhetoric and heated public debates about immigration policies; especially given the challenges associated with distinguishing, at-scale, harmful content from political speech discussing immigration policies.

8.1 Compliance With Meta's Content Policies

The Board concludes that the content does not violate Meta's policies. Thus, Meta's decision to leave the content on Facebook was correct. A minority of the Board believe, however, that Meta's policies could clearly distinguish even the harshest criticisms of immigration policies from speech engaging with conspiracy theories that are harmful toward protected characteristic groups.

I. Content Rules

<u>Hate Speech</u>

The majority of the Board conclude that the content in this case does not violate Meta's Hate Speech Community Standard, and is in fact an example of protected, though controversial, expression of opinion on the topic of immigration. The 50-second clip of Zemmour's interview posted by the user contains no call for violence, nor does it direct dehumanizing or hateful language toward vulnerable groups. The fact that Zemmour has in the past been prosecuted and convicted for use of hateful language, or that the themes of the post bear resemblance to

those of the Great Replacement Theory – which many believe to have sparked violence against migrants and members of minority groups – is not a proper justification for removing a post that does not violate Meta's standards. The policy requires two elements to be present for the content to be considered as violating: (i) a "direct attack" and (ii) a "protected characteristic" group at which the direct attack is aimed. Meta defines "direct attacks" as, among other types of speech, "exclusion or segregation in the form of calls for action," as explained in more detail under Section 4 above. Moreover, the policy rationale makes clear that Meta allows "commentary on and criticism of immigration policies."

For the majority, Zemmour's comments in the video focus mainly on supposed demographical information he presents on Africa, Europe and "colonization." The video contains, among other assertions, the statements, "So the balance of power has reversed," and "When there are now four Africans for one European, what happens? Africa colonizes Europe, and in particular, France." Zemmour's comments do not contain any direct attack, and in fact he does not use the phrase "The Great Replacement" or refer directly to the theory. There is no explicit call to exclude any group from Europe, nor any statement about Africans tantamount to a harmful stereotype, slur or any other direct attack. The Board does, however, find it concerning that Meta does not consider Africans a protected characteristic group given the fact that national origin, race and religion are protected characteristics both under Meta's policies and international human rights law. Africans are mentioned throughout the content. Africa is a collective of nations – thus "Africans" refer to people who are nationals of African countries. Second, in the context of Zemmour's previous comments and discussions about migration in France, the term "Africans" serves as a proxy for non-white Africans, in particular Black and Muslim Africans.

Dangerous Organizations and Individuals

The majority of the Board also conclude that the content does not violate Meta's Dangerous Organizations and Individuals policy, given the lack of elements required to assess this particular piece of content as part of a wider Violence-Inducing Conspiracy Network.

As explained under Section 6 above, Meta considered policy options specific to content discussing conspiracy theories that does not otherwise violate any policies but concluded that, for the time being, implementing any of the options would risk removing a significant amount of political speech. The Board expresses its concern about the lack of information provided by

Meta in response to the Board's questions on this policy development process. The Board notes the company did not provide any specific information about the research it conducted, the information it gathered, the scope of its outreach, types of experts consulted nor the different policy options it analyzed. The Board is also concerned that Meta chose not to share information about the policy development process and its outcome with the public.

A minority of the Board understand that despite the content implicitly targeting several overlapping protected characteristic groups (Black people, Arabs and Muslims), as currently worded, the rules included in Meta's Hate Speech and Dangerous Organizations and Individuals policies do not prohibit content such as this. The fact that the post only repeats the more "palatable" parts of the Great Replacement Theory is, however, not decisive. In the Board's decision on the [Former President Trump's Suspension](#) case, the Board highlighted that Meta "must assess posts by influential users in context according to the way they are likely to be understood, even if their incendiary message is couched in language designed to avoid responsibility."

Nonetheless, as will be explained in more detail in Section 8.3, for a minority of the Board, Meta's approach to content spreading harmful conspiracy theories, such as the Great Replacement Theory, is inconsistent with the aims of the different policies the company has designed to prevent the creation of an environment of intimidation and exclusion that affects protected minorities from online and offline harm. Though a minority of the Board strongly agree that Meta's policies should allow criticism and discussions of all issues (like immigration) that are relevant in democratic societies, they should also establish clear guardrails to prevent the spread of implicit or explicit attacks against vulnerable groups, taking into account the offline harm of certain conspiratorial narratives, such as the Great Replacement Theory.

II. *Transparency*

Meta provides some insight on how it handles immigration-related content in its Transparency Center under the Hate Speech Community Standard in which the company explains that refugees, migrants, immigrants and asylum seekers are protected against the most severe attacks. In a [2017 Newsroom post](#), linked from the Hate Speech policy's rationale, Meta provides some additional detail. However, the company does not explicitly explain how it handles Great Replacement Theory-related content. The 2017 post has information relevant to the topic but was not updated after the 2021 policy development process mentioned under

Section 6 above. Meta also does not explicitly explain in its public-facing policy that calls for exclusion are allowed in the context of discussions on immigration. It is also not clear how implicit or veiled attacks in this context are addressed.

8.2 Compliance With Meta's Human Rights Responsibilities

The majority of the Board find that leaving the content up is consistent with Meta's human rights responsibilities. A minority believe that, in order to be consistent with its human rights responsibilities, Meta needs to reformulate its policies so that its services are not misused by those who promote conspiracy theories that cause online and offline harm.

1. Freedom of Expression (Article 19 ICCPR)

Article 19 of the ICCPR provides for broad protection of the right to freedom of expression, including "freedom to seek, receive and impart information and ideas of all kinds," including "political discourse" and commentary on "public affairs," (General Comment No. 34, para. 11). 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" to protect the rights or reputations of others or to prohibit incitement to discrimination, hostility or violence (General Comment No. 34, para. 11).

In the context of public debates about migration, the UN General Assembly noted its commitment to "protect freedom of expression in accordance with international law, recognizing that an open and free debate contributes to a comprehensive understanding of all aspects of migration." It further committed to "promote an open and evidence-based public discourse on migration and migrants in partnership with all parts of society that generates a more realistic, humane and constructive perception in this regard," (A/RES/73/195, para 33). Immigration and related policies – highly disputed and relevant to political processes not only in France but at a global level – are legitimate topics for debate on Meta's platforms. For the majority, given the potential implications for the public debate, banning this kind of speech on Meta's platforms would be a clear infringement of freedom of expression and a dangerous precedent. For a minority of the Board, it is precisely because open and evidence-based discussions on immigration are so relevant to a democratic society, that the spread of conspiracy theories, such as the Great Replacement Theory, in social media platforms can be

so harmful. [As reported](#) by the Institute for Strategic Dialogue, the methods used to broadcast the theory "include dehumanizing racist memes, distort[ing] and misrepresent[ing] demographic data and us[ing] debunked science."

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. 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).

I.   Legality (Clarity and Accessibility of the Rules)

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

None of Meta's current policies "specifically and clearly" prohibit the content in this case. For the majority of the Board, an ordinary user, reading the Hate Speech Community Standard or Meta's 2017 "Hard Questions" blog post (linked from the Community Standard) would likely get the impression that only the most severe attacks against immigrants and migrants would be removed, as Meta clearly indicates that it wants to allow commentary and criticism of immigration policies on its platforms. The majority of the Board find that this commitment is in line with Meta's human rights responsibilities. For a minority, the Hate Speech policy aims

to prevent the creation of an environment of exclusion or segregation to which hateful conspiracy theories such as the Great Replacement Theory contribute. Given that content engaging with such theories usually targets vulnerable and minority groups and constitutes an attack on their dignity, an ordinary user could expect protection from this type of content under Meta's Hate Speech policy.

Meta's current Dangerous Organization and Individuals policy has no provisions prohibiting the content in this case. For the majority, even if Meta specifically and clearly prohibited content engaging with the Great Replacement Theory on its platforms, the content in this case does not go so far as to name the theory or elaborate on elements of the theory in ways that could be considered conspiratorial and harmful. The post does not allege that migratory flows to Europe involving specific groups of people are part of a secret plot involving actors with hidden agendas.

II. Legitimate Aim

Any restriction on freedom of expression should also pursue at least one of the legitimate aims listed in the ICCPR, which includes protecting the "rights of others." "The term 'rights' includes human rights as recognized in the Covenant and more generally in international human rights law," ([General Comment No. 34](#), para. 28).

In several decisions, the Board has found that Meta's Hate Speech policy, which aims to protect people from harm caused by hate speech, pursues a legitimate aim that is recognized by international human rights law standards (see the [Knin Cartoon decision](#)). It protects the right to life (Article 6, para. 1, ICCPR) as well as the rights to equality and non-discrimination, including based on race, ethnicity and national origin (Article 2, para. 1, ICCPR; Article 2, ICERD). The Board has also previously found that Meta's Dangerous Organizations and Individuals policy seeks to prevent and disrupt real-world harm with the legitimate aim of protecting the rights of others (see the [Shared Al Jazeera Post](#) decision). Conversely, the Board has repeatedly noted that it is not a legitimate aim to restrict expression for the sole purpose of protecting individuals from offense (see [Depiction of Zwarte Piet](#) citing UN Special Rapporteur on freedom of expression, report A/74/486, para. 24, and [Former President Trump's Suspension](#)), as the value that international human rights law places on uninhibited expression is high (General Comment No. 34, para. 38).

III. Necessity and Proportionality

The principle of necessity and proportionality provides that any restrictions on freedom of expression "must be appropriate to achieve their protective function; they must be the least intrusive instrument amongst those which might achieve their protective function; [and] they must be proportionate to the interest to be protected," (General Comment No. 34, para. 34). The nature and range of responses available to a company like Meta are different to those available to a State, and often represent less severe infringements on rights than, for example, criminal penalties. As part of their human rights responsibilities, social media companies should consider a range of possible responses to problematic content beyond deletion to ensure restrictions are narrowly tailored (A/74/486, para. 51).

When analyzing the risks posed by potentially violent content, the Board is guided by the six-part test described in the Rabat Plan of Action, which addresses incitement to discrimination, hostility or violence (OHCHR, A/HRC/22/17/Add.4, 2013). The test considers context, speaker, intent, content and form, extent that the expression has been disseminated and the likelihood of imminent harm.

For the majority of the Board, removal of the content in this case is neither necessary nor proportionate. The Rabat test emphasizes the content and form of speech as "a critical element of incitement." In the content under review in this case, Zemmour's comments, as reproduced in the 50-second clip posted by the user, do not directly engage with the conspiratorial elements of the Great Replacement Theory and the video does not contain inflammatory elements, such as violent or inciting imagery. The comments and the caption also do not contain any direct calls for violence or exclusion. The majority believe it would violate freedom of expression to exclude politically controversial content on the basis of statements made by the speaker elsewhere. The majority view the numbers that Zemmour cites as only slightly exaggerated. The majority also note that the main subject of Zemmour's statements in the video is immigration, perhaps one of today's most salient political issues.

For a minority of Board Members, the content in this case does not violate Meta's current policies (see Section 8.1). However, the company has designed a set of policies aimed at preventing the creation of an environment of exclusion and intimidation that not only affects protected minorities online (impacting the voices of excluded groups) but also offline. Under

these policies, antisemitic and white supremacist narratives, as well as content from Violence-Inducing Conspiracy Networks is moderated. Removing such content is in line with Meta's human rights responsibilities. As explained in Section 2 above, the Great Replacement Theory argues that there is a deliberate plot to achieve the replacement of white populations in Europe with migrant populations predominantly from Africa and Asia. The spread of Great Replacement Theory narratives has contributed to the incitement of racism, hatred and violence targeting immigrants, non-white Europeans and Muslims. A minority of the Board emphasize it is not simply an abstract idea or a controversial opinion but rather a typical conspiracy theory that leads to online and offline harm. It undoubtedly contributes to the creation of an atmosphere of exclusion and intimidation of certain minorities. The evidence of the harm produced by the aggregate or cumulative, scaled and high-speed circulation of antisemitic content on Meta's platforms, as discussed in the [Holocaust Denial](#) case, is similar to the evidence of harm produced by the Great Replacement Theory, indicated under Section 2. For these reasons, a minority find it is inconsistent with the principle of non-discrimination and Meta's values of safety and dignity that Meta has decided to protect certain threatened minority groups from exclusion and discrimination caused by conspiratorial narratives, while keeping others who are in a similar situation of risk unprotected. A minority of the Board found no compelling reason to differentiate Meta's approach to the Great Replacement Theory from the company's approach to other conspiratory narratives mentioned above, which Meta moderates in line with its human rights responsibilities.

Related to the above, for a minority of the Board, the greater challenge faced by social media companies is not in individual pieces of content, but rather in the accumulation of harmful content that is shared on a large scale and at a high speed. The Board has explained that "moderating content to address the cumulative harms of hate speech, even when the expression does not directly incite violence or discrimination can be consistent with Meta's human rights responsibilities in certain circumstances" (see the [Depiction of Zwarte Piet and Communal Violence in the State of Odisha decisions](#)). In 2022, the CERD expressed its concern "at how persistent and widespread racist and discriminatory discourse is [in France], especially in the media and on the Internet." For a minority, the accumulation of Great Replacement Theory-related content "creates an environment where acts of violence are more likely to be tolerated and reproduce discrimination in a society," (see the [Depiction of Zwarte Piet decision](#)). A minority highlights that under the UNGPs "business enterprises should pay special attention to any particular human rights impacts on individuals from groups and populations that may be at a heightened risk of vulnerability and marginalization," (UNGPs Principles 18

and 20). As stated in Section 2 above, the main victims of racism in France are immigrants, especially those of African origin and their descendants. In a [2023 interview](), the Director General of Internal Security for France shared his belief that extremist groups, including those that think they have to take action to stop the "Great Replacement," represent a serious threat in the country.

Even though Meta stated that moderating conspiracy theory-related content would risk removing "an unacceptable amount of political speech," a minority of the Board note the company did not provide any evidence nor data to support that assertion. Moreover, Meta did not explain why this is the case with Great Replacement Theory content but not with, for instance, white supremacist or antisemitic content, since these could also be understood as spreading conspiracy theories. Given the reasons above, for a minority, Meta needs to review its policies to address content that promotes the Great Replacement Theory, unless the company has sufficient evidence: (i) to rule out the harm resulting from the spread of this type of content, as discussed in this decision; or (ii) to demonstrate that the impact of moderating this type of content on protected political speech would be disproportionate. For a proportionate response, among other options, Meta could consider creating an escalation-only policy to allow for the takedown of content openly expressing support of the Great Replacement Theory, without impacting protected political speech, or consider designating actors explicitly engaging with the Great Replacement Theory as part of a Violence-Inducing Conspiracy Network under Meta's Dangerous Organizations and Individuals policy.

The majority is skeptical that any policy could be devised under which this content would be violating, which could satisfy the demands of legality, necessity and proportionality, particularly given the lack of the words "Great Replacement" or any variation thereof in the content. An attempt to remove such content, even taken as a coded reference would result in the removal of significant amounts of protected political expression. Content that is protected on its face should not suffer "guilt by association" either because of the identity of the speaker or the resemblance to hateful ideologies.

9. Oversight Board Decision

The Oversight Board upholds Meta's decision to leave up the content.

10. Recommendations

<u>Transparency</u>

1. Meta should provide greater detail in the language of its Hate Speech Community Standard about how it distinguishes immigration-related discussions from harmful speech targeting people on the basis of their migratory status. This includes explaining how the company handles content spreading hateful conspiracy theories. This is necessary for users to understand how Meta protects political speech on immigration while addressing the potential offline harms of hateful conspiracy theories.

   The Board will consider this implemented when Meta publishes an update explaining how it is approaching immigration debates in the context of the Great Replacement Theory, and links to the update prominently in its Transparency Center.

*Procedural Note:

The Oversight Board's decisions are prepared by panels of five Members and approved by the 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 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. Linguistic expertise was provided by Lionbridge Technologies, LLC, whose specialists are fluent in more than 350 languages and work from 5,000 cities across the world.