

Cases 2024-004-FB-UA, 2024-005-FB-UA, 2024-006-FB-UA – Posts That Include “From the River to the Sea”/ Comments from The Future of Free Speech

1. Background

The Future of Free Speech (FoFS) is an independent, non-partisan think tank located at Vanderbilt University. FoFS’ comment focuses on some of the issues identified by the Oversight Board and particularly Meta’s human rights responsibilities in relation to using the phrase “From the river to the sea”, including freedom of expression, freedom of association, and equality and non-discrimination. This submission will also show how a contextual analysis of the phrase does not justify, from necessity and proportionality perspectives, the introduction of restrictions that would otherwise affect the right to freedom of expression.

2. Key takeaways

- The expression under scrutiny has long been used by various actors around the globe in political debates and protests related to the Israel-Palestine conflict, including the current tensions.
- Any decision regarding the use of such a political slogan must also consider the impact that this may have on the use of this phrase beyond the ongoing situation in the Middle East.
- There is little scope under international human rights law for restrictions on political speech or debates on matters of public interest.
- It is difficult in this case to identify a target or vulnerable group, since expressions fundamentally aim at criticizing the leadership and policies of a state involved in an armed conflict that divides global public opinion.
- The phrase “From the river to the sea” cannot be attributed as such and without any other evidence or context to violent groups or individuals promoting the use of violence.
- Even though this is a phrase that could be potentially used, in certain cases and contexts, to advocate incitement to antisemitic hatred or violence, an assessment of the phrase as such, and without any other contextual element does not meet the Rabat Plan of Action threshold regarding hate speech.
- Only in a context where the phrase is accompanied by explicit references to acts such as ethnic cleansing or endangering the existence of the current state of Israel could it potentially fall under the scope of article 20.2 ICCPR or fulfil the requirements of necessity and proportionality under article 19.3 ICCPR.
- The phrase does not entail a minimally identifiable and specific call to action other than the sensitization of public opinion and the demand for international measures.
- The risk of direct causation of harm is clearly absent.
- Decisions from national authorities in democratic states have acknowledged the importance of considering the phrase in its specific contexts and generally treated it as protected speech, absent further elements that point to antisemitism or incitement to violence.

3. Analysis

3.1. Meta’s human rights responsibilities

3.1.1. *Meta must protect political speech including regarding ongoing international conflicts and international policy debates*

Article 19 of the International Covenant on Civil and Political Rights (ICCPR) recognizes the right to freedom of expression and freedom of information. This provision protects the publication or dissemination by any means of political opinions, ideas, and information, including media editorial content (known as political expression or political speech). This particularly refers to the right not only to express but also to disseminate and receive all sorts of views on all different kinds of political matters, which is the pre-condition for the formation of public opinion and the ultimate expression of political will. In particular, there is little scope under international human rights law for restrictions on political speech or debates on matters of public interest. The UN Human Rights Committee has stressed in this sense that freedom of expression facilitates and reinforces the exercise of other fundamental rights, such as political participation, freedom of assembly and several others.^[1]

The European Court of Human Rights (ECtHR) has also declared that democratic debates do not only require pluralism but also broadmindedness and tolerance. This means that the protection of article 10 of the European Convention on Human Rights (connected to article 19 ICCPR) also applies to those engaging in debates on divisive topics or raising issues of a provocative nature within the context of a certain society and, as it is extensively acknowledged within the case law of the Court, this particularly situates under the safeguards granted to the right to freedom of expression, the rights to “offend, shock, or disturb”.^[2]

The Hamas terrorist attacks of October 7, 2023, and Israel’s military campaign in Gaza have triggered a worldwide wave of debates around very sensitive topics related to the political situation in the Middle East and the role of other States and members of the international community.

Such debates cover areas such as the role of the organization, Hamas, in the governance of Palestine and its relationship with Israel, the responsibility of Hamas in Israel’s response to the attacks, Israel’s policies in the territories of Gaza and the West Bank, strategies to protect and liberate the hostages taken by Hamas, the need to eliminate the capabilities of Hamas and other terrorist groups in the region, allegations of war crimes and genocide committed by Israel in Gaza, the necessity of a humanitarian ceasefire, the role of international organisations, the protection of victims and refugees, the need for the recognition of Palestine as a State, as well as the role and reaction from other Governments and organisations (United States, United Kingdom, European Union, Russia, etc.) to the current situation.

These discussions have also taken place on online platforms, engaging social media users from different parts of the world. In some cases, expressions may reflect frustration, anger or particular claims or aspirations in relation to the conflict. Within the context of social media communication, ideas, opinions and sentiments are often expressed in a simplistic or hyperbolic manner, including via political slogans, memes or hashtags.

As mentioned above, it is not the main aim of this submission to analyse in detail the origin, use and repercussions of the expression “From the river to the sea”. However, it is clear that the posts that generated the cases related to this submission need to be understood within the context of the worldwide online discussions, including on the Facebook platform, regarding the topics mentioned above. In other words, these cannot only be seen as isolated contributions from individual users but as a part of a global conversation where the issue of the status and future of Palestine, as well as its relationship with Israel, is the central theme. In this sense, it is also necessary to underscore that according to the context provided by the Board’s announcement, mentions of both Palestine and Israel need in principle to be understood in their broadest possible meaning: while the former would refer to the territory usually claimed by Palestinian political actors and organisations (and thus not only to Hamas or any other specific actors directly involved in the ongoing situation), the latter needs to be understood as referring to the state of Israel as a political structure (and thus not necessarily to Jewish people or any concrete community present in the country), as the inclusion of hashtags such as #DefundIsrael clearly show.

Finally, it is also important to note that the expression under scrutiny has long been used by various actors around the globe in political debates and protests related to the Israel-Palestine conflict, including the current tensions. This also includes the terrorist group Hamas, whose leaders have also been using it to incite violence since 2012. It is particularly worth mentioning the fact that Israeli politicians have also used the phrase to refer to their aspirations vis-a-vis the same geographical area. In 2020, right-wing lawmaker Gideon Saar stated that “Between the Jordan River and the sea there won’t be another independent state,” meaning a Palestinian state alongside Israel.^[3] Religious Zionist politician Uri Ariel said in 2014, “Between the Jordan River and the Mediterranean Sea there will be only one state, which is Israel.”^[4] Also, a Likud charter from 1977 contains the line: “The right of the Jewish people to the land of Israel is eternal and indisputable... therefore, Judea and Samaria will not be handed to any foreign administration; between the Sea and the Jordan there will only be Israeli sovereignty.”^[5]

It is, therefore, clear that the expression “From the river to the sea” has been used by different political actors and from different perspectives to refer to a specific geographical area and to articulate diverse claims and solutions.^[6] Therefore, any decision regarding the use of such a political slogan must also consider the impact that this may have on the use of this phrase beyond the ongoing situation in the Middle East.

3.1.2. The phrase from the perspective of hate speech policies and violent incitement

The controverted posts have been considered by Meta from the perspective of community standards on hate speech, dangerous organizations and individuals, as well as violence and incitement. Briefly, these standards do not allow direct attacks against people

(rather than concepts or institutions) on the basis of protected characteristics such as race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease. Attacks are to be considered as “violent or dehumanizing speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation”. Also, organizations or individuals that proclaim a violent mission or are engaged in violence (i.e., entities that engage in serious offline harm as well as violent non-state actors that engage in violence against state or military actors in an armed conflict) are not allowed to have a presence on Meta. Finally, Meta also aims to prevent potential offline harm that may be related to content on Facebook. In this context, Meta commits “to consider the language and context in order to distinguish casual statements from content that constitutes a credible threat to public or personal safety”.

Meta must consider a series of important human rights responsibilities when interpreting and enforcing the mentioned content policies.

As already elaborated in a FoFS report, we argue that to be compliant with international human rights standards, a platform’s content moderation practices must be legitimate, necessary, and proportional within the framework of Article 19(3) ICCPR (restrictions on freedom of expression), which sets out the grounds for limitation of freedom of expression.^[7] For hate speech, platforms should frame terms and conditions based on a threshold established by Article 20(2) ICCPR (prohibition of advocacy of hatred) and take strictly into consideration the Rabat Plan of Action’s (RPA) six-part threshold test for context, speaker, intent, content and form, extent of dissemination, and likelihood of imminent harm before taking any enforcement action.^[8] In addition, platforms must refrain from adopting vague blanket policies for removal.

In the case of *Mohamed Rabbae, A.B.S and N.A v The Netherlands* (before the Human Rights Committee),^[9] the HRC gave a relatively extensive analysis of Article 20(2). It held that this article secures the right of persons to be free from hatred and discrimination but holds that it is ‘crafted narrowly’ so as to ensure a protection of free speech. It recalled that free speech may incorporate ‘deeply offensive’ speech, except if the strict threshold of Article 20(2) is met.

It is also important to note, in the specific area of hate speech, that in his report to the UN Human Rights Council of 6 April 2018, the Special Rapporteur on freedom of opinion and freedom of expression acknowledges that entities engaged in content moderation such as Facebook can “regulate” hate speech according to the scale, complexity, and long-term challenges that such form of speech presents on these platforms.^[10] Restrictions could thus be imposed “even if it is not clearly linked to adverse outcomes (as hateful advocacy is connected to incitement in Article 20(2) of the ICCPR)”. In any case, companies should articulate the bases for such restrictions, and demonstrate the necessity and proportionality of their actions. Based precisely on this, the Board considered in the case “Armenians in Azerbaijan” (use of the word “tazis” to refer to Azerbaijanis) that the slur used was “hateful and dehumanizing”. Although the expression could not be qualified as incitement, “the potential for adverse outcomes was nevertheless present”. The use of dehumanizing language in a context of armed conflict may create “a discriminatory environment that undermines the freedom of others to express themselves”. Therefore, the Board concluded that “the presence of these risks and Facebook’s human rights responsibility to avoid contributing to them meant it was permitted to remove the slur.”

The RPA states that there must be a high threshold when applying Article 20 of the ICCPR. Further, as noted in the 2012 Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, ‘the threshold of the types of expression that would fall under the provisions of Article 20(2) should be high and solid.’^[11] To achieve this threshold, the RPA sets out a six-part threshold test to be referred to when applying Article 20(2) and includes the assessment of the (i) social and political context (ii) status of the speaker, (iii) intent to incite the audience against a target group (iv) content and form of the speech (v) extent of its dissemination and (vi) likelihood of harm, including imminence. Also, the Council of Europe elaborated, based on the RPA, a toolkit for analyzing a case of hate speech which can also provide very useful guidance to navigate the meaning and implications of each criterion.^[12]

Social and political context

Regarding the social and political context, it is important that the Board makes a proper assessment on whether the statements are likely and deliberately posted to incite discrimination, hostility or violence against a specific target group. Analysis of the context should place the speech act within the social and political context that has already been described above, where different online voices are expressing the condemnation of a series of military operations as well as pointing at a desired outcome (the establishment of Palestine as an independent state). This analysis must also acknowledge the difficulties in this case to establish the existence of

such a target or vulnerable group (in the sense of a minority from an ethnic / racial / religious / sexual / gender orientation / social status / other criteria view and with lack of a position of equal power), since the expressions fundamentally aim to criticize the leadership and policies of a state involved in an armed conflict dividing global opinion.

Status of the speaker

Regarding the status of the speaker, it must be noted that in the cases considered by the Board we are not in the presence (either directly or indirectly) of organizations or individuals proclaiming a violent mission or engaging in violence. As already mentioned, the phrase “From the river to the sea” cannot be exclusively attributed as such and without any other evidence to violent groups or individuals promoting the use of violence. More particularly, such expression cannot be considered per se as an endorsement or expression of sympathy vis-à-vis the terrorist organization Hamas or any other violent extremist group. Since it has been reproduced, chanted and disseminated by students, peaceful demonstrators, political organizations and many other non-violent actors, no immediate conclusions in terms of hate speech can be established from the perspective of the speaker. This is also what can be inferred from the description provided by the Board in the announcement and brief description of the cases.

Intent

Article 20.2 of the ICCPR anticipates intent. According to the RPA, negligence and recklessness are not sufficient for an act to be an offence under article 20.2 of the ICCPR, as this article provides for “advocacy” and “incitement” rather than the mere distribution or circulation of material. In this regard, it requires the activation of a triangular relationship between the object and subject of the speech act as well as the audience. Even though, and as already mentioned, the UN Special Rapporteur on freedom of opinion and freedom of expression has established specific recommendations applicable to hate speech on online platforms in the sense that actual incitement is not absolutely required, it is also true that Meta’s community standards require, in line with international human rights law, the existence of “violent or dehumanising speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation”. The phrase in question, considered as such, only expresses, in a simplified manner, an opinion and an aspiration. In other words, only in a context where the phrase is accompanied by explicit references to acts of ethnic cleansing or endangering the existence of the state of Israel, could it potentially fall under the scope of article 20.2 ICCPR, or fulfil the requirements of necessity and proportionality under article 19.3 ICCPR. It is also important to underscore the fact that in terms of intent, the phrase can be generally associated with a privileged form of speech: the dissemination of political ideas and opinions, including in the international public sphere.

Content and form

The latter is connected to the criterion of content and form, where it is important to note that apart from an essentially political expression, the phrase does not entail a minimally identifiable and specific call to action, other than to influence public opinion and reinforce the demand for measures to be adopted by actors (humanitarian organizations, international organizations, national governments and institutions, etc.) with the capacity to intervene in a non-violent manner. In other words, and once again, in the absence of more explicit messages or calls for specific action, the phrase cannot, as such, be necessarily connected to hate narratives that would make the introduction of restrictions legitimate and necessary.

Extent of the act

In terms of the extent of the act, dissemination via social media of these messages makes their potential audience very wide. However, the Board shall also consider the fact that this kind of short message/hashtag is very often disseminated by individual accounts with diverging (and not necessarily high) numbers of followers. Only cases of combined social preeminence of the speaker, significant audience and presence of additional and inciting content would thus deserve special consideration in terms of the possible existence of hate speech.

Likelihood of harm

Finally, in terms of the likelihood of harm, including imminence, the RPA clearly establishes that some degree of risk of harm must be identified. It means that the competent bodies (in this case the Board) will have to determine that there was a reasonable

probability that the speech would succeed in inciting actual action against the target group, recognizing that such causation should be rather direct. All the considerations already formulated in this submission regarding other criteria suggest that, in this case, the mentioned direct causation is absent.

3.1.3. Decisions by national authorities

Various national authorities in democratic states have adopted different approaches to the use of the phrase, even though the number of elaborated court decisions on this matter is very limited. Most judicial reactions are also very much so context-specific.

In the United Kingdom, London's Metropolitan Police said it was legal when activists projected the phrase onto the side of the Houses of Parliament in February 2024.^[13] However, they do caveat their statement by saying, "While there are scenarios where chanting or using these words could be unlawful depending on the specific location or context".^[14] Previously, in October 2023 the Met Police also said they would not be arresting anyone chanting the slogan used by Palestinian activists during demonstrations.^[15] The caveat again was added: while the use of the chant in order to intimidate or harass Jewish people would be an offence, "it is likely that its use in a wider protest setting, such as we anticipate this weekend, would not be an offence and would not result in arrests".^[16]

In Germany, the Ministry of Justice tweeted a statement establishing that: "'From the river to the sea' is a Hamas slogan. Hamas is prohibited from operating in Germany. Depending on the circumstances of the individual case, the use of the slogan can, therefore, be considered the use of propaganda by a banned organization. This is a criminal offense."^[17] Even though the slogan is connected in this case to a terrorist organization, the authorities also open the door to a proper consideration of "the circumstances of the individual case". In addition to this, in December 2023 two German courts analyzed this slogan (together with other ones) and established that criticism of the Israeli state is "protected by freedom of expression", stressing that the slogans are "directed against Israel and not against the Jewish population of Germany", which resulted in the rejection of "inflammatory allegations of antisemitism".^[18] A similar decision was recently adopted by a court in the Netherlands.^[19]

However, in October 2023 Austrian police banned a pro-Palestinian protest in Vienna due to coincide with a pro-Israel event after the (at the time) very recent attack by Hamas, citing the fact the phrase "From the river to the sea" was mentioned in invitations to the demonstration and characterising it as a call to violence.^[20] It should be noted, in any case, that this happened at the very early stage of Israel's response in Gaza and when the images of Hamas' terrorist acts were particularly present in the mindset of policymakers and the whole population. Special protections and limitations regarding possible antisemitic speech are also present in the country due to obvious historical reasons.

Conclusion

In the context of conflict, words and expressions may cause serious harm and constitute serious incitement. In addition to this, expressions targeting a special group or minority, which may entail any form of dehumanization or calls for persecution, need to be scrutinized with special care. At the same time, social media platforms have also become important actors in the facilitation of global discussions around international policy and humanitarian implications of certain conflicts, thus facilitating better scrutiny of the position or actions of states and the international community.

It is, therefore, necessary for social media platforms like Facebook or Instagram to adhere to international human rights law to guarantee a proper assessment of certain expressions. This human rights framework enables platforms properly to protect the dissemination of politically engaged opinions while also limiting, when strictly necessary and justified, messages that contain a clear and unequivocal incitement, based on the criteria presented in this submission.

In particular, the phrase "From the river to the sea" has long been used by various actors around the globe in political debates and protests related to the Israel-Palestine conflict, including the current tensions. This is a phrase that indeed could be potentially used, in certain cases and contexts, to advocate incitement to antisemitic hatred or violence. However, an assessment of the phrase as such, and without any other contextual element with a clear incitement component, does not meet the Rabat Plan of Action threshold regarding hate speech.

[1] General Comment no. 34. CCPR/C/GC/34, paragraph 4.

[2] *Handyside v. The United Kingdom*. Application no. 5493/72. Decision of 7 December 1976. Paragraph 49. See also Council of Europe, “Freedom of political speech: an imperative for democracy”, Information Documents SG/Inf(2022)36, 6 October 2022. Available at <https://rm.coe.int/0900001680a86af5>

[3] <https://www.ynet.co.il/news/article/5643203>

[4] <https://www.srugim.co.il/68778-תהי-לים-הירדן-בין-אריאל-בין-הירדן-לימים-תהי-שר-השיכון-אורי-אריאל-בין-הירדן-לימים-תהי-68778>

[5] https://www.jewishvirtuallibrary.org/original-party-platform-of-the-likud-party?utm_content=cmp-true

[6] See <https://apnews.com/article/river-sea-israel-gaza-hamas-protests-d7abbd756f481fe50b6fa5c0b907cd49#> From a broader perspective see also Joe Hernandez, “How interpretations of the phrase ‘from the river to the sea’ made it so divisive”, *NPR* November 9 2023, available at <https://www.npr.org/2023/11/09/1211671117/how-interpretations-of-the-phrase-from-the-river-to-the-sea-made-it-so-divisive> In addition to this, according to a recent University of Chicago poll, when Muslim students hear “From the River to the Sea,” 76% understand it to mean Arabs and Jews living together in two states or one state; but 66% of Jewish students take it to mean Palestinians should replace Israelis in the territory (quoted here: <https://x.com/shibleyvelhami/status/1790903654829985802?s=46>). Finally, in this interesting podcast Deborah Brown of Human Rights Watch and Evelyn Douek of Stanford Law talk with Jameel Jaffer about the role that social media platforms are playing in shaping, suppressing, and distorting public discourse about the war (<https://podcasts.google.com/feed/aHR0cHM6Ly9mZWVkcyc5idXp6c3Byb3V0LmNvbS8yMDg3ODMxLnJzcw/episode/OnV6enNwcm91dC0xNTA3NjI2NA?ep=14>).

[7] “A framework of first reference. Decoding a human rights approach to content moderation in the era of ‘platformization’”. Available at https://futurefreespeech.org/wp-content/uploads/2021/11/Report_A-framework-of-first-reference.pdf

[8] Available at <https://www.ohchr.org/en/freedom-of-expression>

[9] CCPR/C/117/D/2124/2011. Available at <https://juris.ohchr.org/casedetails/2153/en-US>

[10] A/HRC/38/35. Available at https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/35

[11] A/67/357. Available at https://ap.ohchr.org/documents/dpage_e.aspx?si=A/67/357

[12] Available at <https://rm.coe.int/advanced-guide-toolkit-how-to-analyse-hate-speech/1680a217cd>

[13] <https://x.com/metpoliceuk/status/1760631175272349861>

[14] <https://x.com/metpoliceuk/status/1760631175272349861>

[15] <https://www.middleeasteye.net/news/israel-palestine-war-uk-police-river-sea-chant-not-arrestable>

[16] <https://www.middleeasteye.net/news/israel-palestine-war-uk-police-river-sea-chant-not-arrestable>

[17] https://twitter.com/bmj_bund/status/1789963550024163589

[18] <https://www.newarab.com/news/germany-two-courts-say-pro-palestinian-slogans-legal>

[19] <https://novamedia.com/2023/10/18/dutch-court-rules-from-the-river-to-the-sea-protected-speech-and-not-antisemitic/>

[20] <https://www.reuters.com/world/from-river-sea-prompts-vienna-ban-pro-palestinian-protest-2023-10-11/>