Hate Speech Rules Under International Law

Toby Mendel
Executive Director
Centre for Law and Democracy

February 2010

I. Introduction

The issue of what speech should or may be prohibited on the basis that it incites others to hatred – so-called 'hate speech' – is a matter of great dispute and argumentation globally, although the standards on this under international law are in fact reasonably developed. International law not only allows, but actually requires, States to ban certain speech on the basis that it undermines the right of others to equality or to freedom from discrimination, and occasionally also on the basis that this is necessary to protect public order.

Promoting substantive equality among human beings, including freedom from discrimination, is a foundational idea in human rights, and this is reflected in the very first article of the *Universal Declaration on Human Rights* (UDHR), adopted by the UN General Assembly in 1948,[1] which states: "All human beings are born free and equal in dignity and rights." The second article of the UDHR provides for equal enjoyment of the rights and freedoms proclaimed, "without distinction of any kind, such as race, colour, sex, …" and several other articles refer explicitly to the equal enjoyment of various rights.[2]

This chapter reviews the international law standards regarding hate speech, and how they have been interpreted by international courts.

II. International Provisions Freedom of Expression and Hate Speech

The right to freedom of expression is a fundamental human right which finds protection in all major human rights systems, as well as in national constitutions. At the same time, it is not an absolute right, and it may be limited to protect overriding public and private interests, including equality and public order. International law contains a number of provisions which provide a framework for balancing freedom of expression against these other interests in the particular context of hate speech.

Article 19 of the UDHR guarantees the right to freedom of expression, including to "seek, receive and impart information and ideas through any media and regardless of frontiers."

---

[1] General Assembly Resolution 217A(III), 10 December 1948.

[2] Article 7 provides for equality before the law, Article 10 for equality in public hearings and Article 21(2) for equal access to public service.

Perhaps surprisingly, given that it reflects, at least in part, a global reaction to the excesses of the Second World War, the UDHR does not specifically provide for prohibitions on hate speech or incitement to hatred. Article 7 does, however, provide for equal protection for all against discrimination in violation of the Declaration, and also against incitement to discrimination.

The first international treaty to deal directly with the issue of hate speech was the *International Convention on the Elimination of all Forms of Racial Discrimination* (CERD), adopted by the UN General Assembly in 1965.[3] Its provisions are not only the first to address hate speech, but also by far the most far-reaching.

It is probably useful to distinguish four different aspects of the hate speech obligations provided for in CERD, found in its Article 4(a):[4]

1. dissemination of ideas based on racial superiority;
2. dissemination of ideas based on racial hatred;
3. incitement to racial discrimination; and
4. incitement to acts of racially motivated violence.

Article 4(c) also calls for prohibitions on public authorities or institutions promoting or inciting racial discrimination, illustrating the particular evil of public officials and bodies engaging in racist activities.

CERD, by virtue of its focus on racial discrimination, does not guarantee the right to freedom of expression. However, Article 4 does require that any measures taken to implement it have due regard for the principles set out in both the UDHR and Article 5 of CERD, which provides for equality before the law in the enjoyment of a large number of rights, including freedom of expression.[5]

Article 19(2) of the *International Covenant on Civil and Political Rights* (ICCPR), adopted by the UN General Assembly in 1966,[6] guarantees the right to freedom of expression as follows:

> Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art or through any other media of his choice.

Article 19(3) of the ICCPR permits limited restrictions on freedom of expression where these are a) provided by law; b) for the protection of one of the legitimate interests listed; and c) necessary to protect that interest.

---

[3] General Assembly Resolution 2106A(XX), 21 December 1965, entered into force 4 January 1969. There were 173 Parties and six additional signatories to CERD as of 11 February 2010.
[4] In its General Comment No. 15 of 23 March 1993, the CERD Committee refers to four categories.
[5] Article 5(d)(viii).
[6] General Assembly Resolution 2200A (XXI), 16 December 1966, entered into force 3 January 1976. There were 165 Parties and seven additional signatories to the ICCPR as of 11 February 2010.

The ICCPR places an obligation on States Parties to prohibit hate speech in rather different terms than CERD. Article 20(2) provides:

> Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.

All of the three regional human rights treaties – the *European Convention on Human Rights* (ECHR),[7] the *American Convention on Human Rights* (ACHR)[8] and the *African Charter on Human and Peoples' Rights* (ACHPR)[9] – guarantee the right to freedom of expression, respectively at Article 10, Article 9 and Article 13. These guarantees are largely similar to those found in the ICCPR.

Perhaps surprisingly, only the ACHR specifically provides for the banning of hate speech, at Article 13(5), as follows:

> Any propaganda for war and any advocacy of national, racial, or religious hatred that constitute incitements to lawless violence or to any other similar illegal action against any person or group of persons on any grounds including those of race, color, religion, language, or national origin shall be considered as offenses punishable by law.

III. The Relationship Between Articles 19 and 20 of the ICCPR

Articles 19(3) and 20(2) of the ICCPR are clearly in potential tension and even conflict with each other, and for this reason they received a lot of attention during the drafting process.[10] They were, quite properly, kept separate since Article 19 guarantees freedom of expression, while permitting restrictions on that right, whereas Article 20(2) imposes an obligation to restrict speech. However, it was decided that they should go next to each other, to emphasise the close relationship between them.[11]

The UN Human Rights Committee (HRC), the body of experts tasked with interpreting the ICCPR, has specifically stated that Article 20(2) is compatible with Article 19.[12] As a result, any law seeking to implement the provisions of Article 20(2) must not overstep the permissible scope of restrictions on freedom of expression allowed by Article 19(3). Conversely, Article 19(3) must be interpreted in a manner that respects the terms of Article 20(2). This is supported by the case law. In the case of *Ross v. Canada*, for example, the HRC specifically held that a

[7] Adopted 4 November 1950, entered into force 3 September 1953.
[8] Adopted 22 November 1969, entered into force 18 July 1978.
[9] Adopted 26 June 1981, entered into force 21 October 1986.
[10] See Bossuyt, M., *Guide to the "Travaux Préparatoires" of the International Covenant on Civil and Political Rights* (Dordrecht: Martinus Nijhoff Publishers, 1987), pp. 398-411.
[11] *Ibid.*, p. 406.
[12] General Comment 11: Prohibition of propaganda for war and inciting national, racial or religious hatred (Art. 20), 29 July 1983.

restriction on racist expression had to be justified by reference to the test set out in Article 19(3) of the ICCPR.[13]

An interesting question is whether Article 19(3) would permit restrictions on hate speech beyond the scope of what Article 20(2) requires. Theoretically, this is possible: what States are required to ban to ensure equality is not necessarily the same as what they are permitted to ban to serve this goal without breaching the right to freedom of expression. At the same time, the drafting history of Article 20(2) suggests that there was little scope for extending its provisions while still respecting Article 19. Proposals to restrict Article 20(2) to incitement to violence were rejected, but so were proposals to extend it, for example to include 'racial exclusiveness', on the basis of concern about free speech.[14] This suggests that the obligations of Article 20(2) are either identical or extremely close to the permissions of 19(3).

Faurisson v. France, a case in which the HRC was called upon to assess a conviction under a law which prohibited any contestation of the existence of the category of crimes against humanity defined in the Nuremburg Charter, sheds important light on this issue. Evatt, Kretzmer and Klein, in a concurring opinion, stated:

> [T]here may be circumstances in which the right of a person to be free from incitement to discrimination on grounds of race, religion or national origins cannot be fully protected by a narrow, explicit law on incitement that falls precisely within the boundaries of article 20, paragraph 2. This is the case where … statements that do not meet the strict legal criteria of incitement can be shown to constitute part of a pattern of incitement against a given racial, religious or national group, or where those interested in spreading hostility and hatred adopt sophisticated forms of speech that are not punishable under the law against racial incitement, even though their effect may be as pernicious as explicit incitement, if not more so.[15]

This can be understood as a call for an extremely narrow and precise interpretation of incitement in Article 20(2), alongside a recognition that there may be special cases where statements which do not fall within the scope of this very narrow interpretation may still legitimately be prohibited because, in context and alongside other statements, they in fact constitute a pattern of incitement. Alternately, their point could be understood as a comment on the proper interpretation of incitement, rather than going outside of the boundaries of Article 20(2), *per se*. Either way, the analysis confirms that Article 19(3) and Article 20(2) are legally contiguous or very nearly so.

IV. Key Elements of the Offence of Hate Speech

International law, and specifically the right to freedom of expression, imposes constraints on what may be banned as hate speech. Three key aspects of hate speech

[13] 18 October 2000, Communication No. 736/1997, para. 11.1.
[14] Bossuyt, note 10, pp. 404-405, 408.
[15] 8 November 1986, Communication No. 550/1993, para. 4.

– namely intent, incitement and what results are prohibited – are examined in this section.

III.1 Intent

Article 20(2) of the ICCPR and Article 13(5) of the ACHR require advocacy of hatred, while Article 4(a) of CERD does not. The advocacy element can be understood as an intent requirement, so that only statements made with the intent of inciting hatred are covered.

The importance of including intent in hate speech crimes is illustrated well by a case from Denmark. In that case, Jersild, a journalist, had been convicted for a television programme which included hate speech statements by racist extremists, although the purpose of the programme was really to expose racism in Denmark. The Committee on the Elimination of all Forms of Racial Discrimination (CERD Committee) was divided in its response, stating, in its report to the UN General Assembly:

> Some members welcomed this decision as the clearest statement yet, in any country, that the right to protection against racial discrimination took precedence over the right to freedom of expression. Other members thought that in such cases the facts needed to be considered in relation to both rights.[16]

The European Court of Human Rights, by a clear majority but not a unanimous decision, held that the conviction by the Danish courts was a breach of Jersild's right to freedom of expression.[17] The Court took into account the fact that the statements were made in the context of a serious programme intended for an informed audience and dealing with social and political issues. It also relied heavily on its finding that Jersild's purpose or intent was not to promote racism but, on the contrary, to expose and analyse it, stating:

> [A]n important factor in the Court's evaluation will be whether the item in question, when considered as a whole, appeared from an objective point of view to have had as its purpose the propagation of racist views and ideas.[18]

In a concurring opinion in *Faurisson*, Evatt, Kretzmer and Klein expressed concern that the law pursuant to which the author of the complaint was convicted did "not link liability to the intent of the author".[19] However, based on the facts, the Committee decided that Faurisson had clearly been motivated by a desire to

[16] Report of the Committee to the General Assembly, Official Records, Forty-Fifth Session, Supplement No. 18 (A/45/18), p. 21, para. 56..
[17] *Jersild v. Denmark*, 22 August 1994, Application No. 15890/89. The Court noted the difference of opinion within the CERD Committee and also opined that its decision was compatible with CERD. See paras. 21 and 30.
[18] *Ibid.*, para. 31.
[19] Note 15, para. 9.

promote racism and, as a result, the conviction in that particular case was legitimate.[20]

In light of these cases, the lack of an intent requirement in CERD presents a clear and problematical conflict with the right to freedom of expression.

III.2 Incitement

Article 7 of the UDHR, Article 20(2) of the ICCPR and Article 13(5) of the ACHR apply only where incitement is present. Two of the four relevant provisions in Article 4(a) of CERD require incitement, but the other two prohibit the mere dissemination of certain ideas, namely those based on superiority and racial hatred. This was a matter of some controversy when Article 4 of CERD was discussed at the UN General Assembly, and a motion to delete these provisions was tabled but defeated.[21]

The question of what constitutes incitement is extremely complex and controversial; the UN High Commissioner for Human Rights, for example, has expressed concern at the fact that this term lacks clear definition in international law.[22] International courts have looked at a number of factors when assessing whether incitement is present, focusing on the nexus between the statements and the proscribed result, and issues such as causation and context.

<u>Causation</u>

It is clear that inciting an act is not the same thing as causing it. At the same time, international courts often look for causation-related factors when assessing whether speech incites hatred. In the case of *Ross v. Canada*, a teacher was removed from the classroom for his anti-Semitic/Holocaust denial publications. The Supreme Court of Canada noted the evidence that a 'poisoned environment' had been created within the relevant school board and held that "it is possible to 'reasonably anticipate' the causal relationship" between that environment and the author's publications.[23] The HRC held that this satisfied the necessity part of the test for restrictions on freedom of expression and that, as a result, there was no breach of this right.[24]

In at least one case from Turkey involving allegations of hate speech, *Erbakan v. Turkey*, the European Court of Human Rights found a breach of the right to freedom

[20] *Ibid.*, para. 10. See also the concurring decision of Lallah, paras. 6 and 9.
[21] See Lerner, N., *The U.N. Convention on the Elimination of all Forms of Racial Discrimination* (Alphen aan den Rijn, The Netherlands: Sijthoff & Noordhoff, 1980), p. 46. The vote on deleting them was 54 against, 25 in favour and 23 abstentions.
[22] Study of the United Nations High Commissioner for Human Rights compiling existing legislations and jurisprudence concerning defamation of and contempt for religions, UN Doc. A/HRC/9/25, 5 September 2008, para. 24.
[23] *Ross v. New Brunswick School District No. 15*, [1996] 1 S.C.R. 825, para. 101.
[24] Note 13, para. 11.6.

of expression on the basis that the impugned statements did not create an actual risk of harm, stating:

> [I]t was not established that at the time of the prosecution of the applicant, the impugned statements created an "actual risk" and an "imminent" danger for society ... or that they were likely to do so.[25]

Often, however, as the International Criminal Tribunal for Rwanda (ICTR) noted in its assessment of international hate speech cases in the *Nahimana* case, international courts do not look at the matter from a direct causal perspective: "Rather, the question considered is what the likely impact might be, recognizing that causation in this context might be relatively indirect."[26] Thus, in the *Faurisson* case, the HRC noted that the impugned statements, "were of a nature as to raise or strengthen anti-Semitic feelings".[27]

A series of claims that hate speech convictions represented a breach of the right to freedom of expression, which were rejected by the European Commission and Court of Human Rights as inadmissible, focused on impact. Little reasoning was provided in most of these cases to substantiate the claimed impact; instead, the Commission and Court quickly concluded that where the statements in question were racist, they were likely to undermine other rights, in particular equality.[28] In some cases, the Commission or Court referred to the likelihood of the impugned statements fostering anti-Semitism.[29] In others, the negative impact of the statements on the underlying Convention objectives of justice and peace was noted.[30]

The causality or likely impact standards employed in these cases are weak, a problem which is exacerbated by the very general nature of the aims protected, such as justice, peace or freedom from hatred. Thus, in *Ross*, the standard was 'possible to reasonably anticipate' and in *Faurisson*, 'of a nature to raise', while in the other cases no specific standard was even mentioned.

Context

[25] *Erbakan v. Turkey*, 6 July 2006, Application No. 59405/00, para. 68. Unofficial translation from the original French: "il n'est pas établi qu'au moment de l'engagement des poursuites à l'encontre du requérant, le discours incriminé engendrait « un risque actuel » et un danger « imminent » pour la société (paragraphe 48 ci-dessus) ou il était susceptible de l'être.
[26] *Prosecutor v. Nahimana, Barayagwiz and Ngeze*, 3 December 2003, ICTR-99-52-T (Trial Chamber), para. 1007.
[27] Note 15, para. 9.6.
[28] See *Glimmerveen and Hagenbeek v. Netherlands*, 11 October 1979, Application No. 8406/78; *Kühnen v. Germany*, 12 May 1988, Application No. 12194/86; *B.H., M.W., H.P. and G.K. v. Austria*, 12 October 1989, Application No. 12774/87 (*B.H., M.W., H.P. and G.K.*); *Ochensberger v. Austria*, 2 September 1994, Application No. 21318/93; *Nationaldemokratische Partei Deutschlands, Bezirksverband München-Oberbayern v. Germany*, 29 November 1995, Application No. 25992/94; *Remer v. Germany*, 6 September 1995, Application No. 25096/94; and *Garaudy v. France*, 7 July 2003, Application No. 65831/01.
[29] See *Kühnen* and *Garaudy*.
[30] See *Remer*, *Nationaldemokratische Partei Deutschlands* and *Garaudy*

Context is clearly of the greatest importance in assessing whether particular statements are likely to incite to hatred – as it may have a bearing on both intent and causation – and many of the hate speech cases refer to contextual factors. In *Faurisson*, for example, the HRC noted a statement by the, "then Minister of Justice, which characterized the denial of the existence of the Holocaust as the principal vehicle for anti-Semitism." The concurring opinion by Evatt, Kretzmer and Klein also referred to this problem, stating:

> The notion that in the conditions of present-day France, Holocaust denial may constitute a form of incitement to anti-semitism cannot be dismissed. This is a consequence not of the mere challenge to well-documented historical facts, established both by historians of different persuasions and backgrounds as well as by international and domestic tribunals, but of the context, in which it is implied, under the guise of impartial academic research, that the victims of Nazism were guilty of dishonest fabrication, that the story of their victimization is a myth and that the gas chambers in which so many people were murdered are "magic".[31]

Similarly, in *Ross*, the HRC, in line with decisions at the national level, was very sensitive to the contextual fact that the author had been a teacher:

> In the circumstances, the Committee recalls that the exercise of the right to freedom of expression carries with it special duties and responsibilities. These special duties and responsibilities are of particular relevance within the school system, especially with regard to the teaching of young students.[32]

In two cases from Turkey, the European Court distinguished otherwise arguably similar situations on the basis of context.[33] In *Zana*, the Court refused to find a breach of the right to freedom of expression for a conviction in Turkey on the basis that the applicant "defended an act punishable by law as a serious crime" and "endangering public safety".[34] The Court took into account contextual factors such as the fact that the applicant was a former major of a town in south-east Turkey and that the statements "coincided with murderous attacks" in the area.[35] In *Incal*, the Court found a breach of the right to freedom of expression, stating that, "the circumstances of the present case are not comparable to those found in the *Zana* case. Here the Court does not discern anything which would warrant the conclusion that Mr Incal was in any way responsible for the problems of terrorism in Turkey, and more specifically in İzmir."[36] [references omitted]

III.3 Proscribed Results

Different international rules call for the prohibition of statements inciting different proscribed results. Article 13(5) of the ACHR is limited to incitement to violence or

[31] Note 15, para. 6.
[32] Note 13, para. 11.6.
[33] The State, at least, argued that the cases were similar. See *Incal v. Turkey*, 9 June 1998, Application No. 22678/93, para. 44.
[34] *Zana v. Turkey*, 25 November 1997, Application No. 18954/91, para. 26.
[35] *Ibid.*, para. 59. See generally paras. 58-60.
[36] Note 33, para. 58.

similar illegal actions. Both Article 4(a) of CERD and Article 20(2) of the ICCPR go beyond that to additionally cover incitement to discrimination and hatred (or hostility). Article 4(a) of CERD goes even further, calling for the prohibition of all ideas based on superiority.

Two of these results – violence and discrimination – are (normally) illegal acts and thus these rules are simply a species of a more general rule prohibiting incitement to crime. Hatred, on the other hand, as understood by both the HRC and the CERD Committee, is a state of mind rather than a specific act. Furthermore, hatred, as such, is simply an opinion and is thus absolutely protected under international law. Despite this, and although there are some States which persistently object to the idea of banning incitement to hatred,[37] most States accept this international law standard. This is at least partly because hatred will inevitably find some form of tangible manifestation, and groups should not have to wait until concrete acts are perpetrated on them before being able to claim some protection.

Perhaps surprisingly, there is a distinct paucity of material in official documents and court decisions on what constitutes 'hatred'.[38] The Council of Europe Recommendation on Hate Speech defines 'hate' implicitly, as follows:

> [T]he term "hate speech" shall be understood as covering all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, antisemitism or other forms of hatred based on intolerance, including: intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants and people of immigrant origin.[39]

This definition is largely circular, although it does provide a few notions that can help anchor the meaning of the term hate, namely 'xenophobia', 'anti-Semitism', 'intolerance', 'aggressive nationalism', 'ethnocentrism' and 'discrimination'.

In the *Nahimana* case, the ICTR defined hate speech as, "stereotyping of ethnicity combined with its denigration".[40] However, both 'stereotyping' and 'denigration' are fairly elastic terms, which, furthermore, impose a rather low standard. Much stereotyping, both positive and negative, is perfectly legitimate, or at least certainly should not be banned, as manifested, for example, in the comedy routines that are common in many countries.

In the hate speech cases before the HRC and European Commission and Court, the focus tends to be more on whether the statements in question are of a racist nature,

[37] Notably the United States.
[38] The UN High Commissioner for Human Rights has expressed concern about this. See note 22.
[39] Appendix, under Scope.
[40] Note 26, para. 1021.

or perhaps on the harm they may have on the rights of others, rather than on defining hate as such. [41]

There have been numerous academic attempts to distinguish hate speech from merely offensive speech. One line of reasoning, which is helpful at least conceptually, is to distinguish between expression targeting ideas, including offensive expression, which is protected, and abusive expression which targets human beings, which may not be protected.[42] In *Giniewski v. France*, the European Court of Human Rights seemed to support this approach, holding that the impugned speech was not a gratuitous attack on religion but, rather, part of a clash of ideas ('débat d'idées').[43]

V. Conclusion

There are important differences in terms of both scope and balancing of rights between the different international law regimes governing hate speech. This has created problems for international courts and other bodies tasked with interpreting human rights standards, and has lead to confusion as to what, exactly, is hate speech. Even where the legal rules are reasonably precise and coherent, as it the case under the ICCPR, interpretation has been problematical. This is particularly true of 'incitement', which is the touchstone of what constitutes hate speech.

Despite this, some general conclusions may be drawn. Statements which are made without an intention to incite others to hatred should not be proscribed as hate speech. To constitute incitement, there must be a close nexus between the relevant statement and the risk of harm, for example as evidenced by causality or likely impact. However, specific (and illegal) acts such as discrimination and violence are not the only recognised harms; it is enough if statements incite others to hatred, even if they do not act on it.

[41] See Oetheimer, M., "Protecting Freedom of Expression: The Challenge of Hate Speech in the European Court of Human Rights Case Law" (2009) 17 *Cardozo Journal of International and Comparative Law* 427, p. 429.
[42] See, for example, Gaudreault-DesBiens, J., "From Sisyphus's Dilemma to Sisyphus's Duty? A Meditation on the Regulation of Hate Propaganda in Relation to Hate Crimes and Genocide" (2000) 46 *McGill Law Journal* 121, p. 135.
[43] 31 January 2006, Application No. 64016/00, para. 50.