



THE HIGH LEVEL PANEL OF LEGAL EXPERTS ON MEDIA FREEDOM

The independent advisory body to the Media Freedom Coalition



Human Rights
Institute



Hate Speech in International Law

Authors: Ms Amal Clooney and Ms Alice Gardoll

Excerpt from *Freedom of Speech in International Law* (edited by Lord David Neuberger and Ms Amal Clooney, published by Oxford University Press).



This IBAHRI publication is an extracted chapter of Freedom of Speech in International Law, edited by Lord David Neuberger of Abbotsbury and Ms Amal Clooney and published by Oxford University Press in January 2024. The authors of this chapter are Ms Amal Clooney and Ms Alice Gardoll. Ms Alice Gardoll is the Assistant Editor of the text. The full text can be purchased in hard copy or as an e-book [here](#). All sources and references have been checked for accuracy as of 1 January 2022.

Freedom of Speech in International Law outlines the minimum protections for speech enshrined in international law, focusing on four types of laws that are being weaponized to silence the press and independent voices: laws regulating defamatory or insulting speech, laws regulating false speech, laws regulating hate speech and laws regulating national security. The book provides examples of where states are falling short and makes recommendations about how international standards should be interpreted, updated and enforced.

Recommendations are based on international legal standards that apply to states and that many social media companies have expressed adherence to. The recommendations have been endorsed by the High Level Panel of Legal Experts on Media Freedom, as well as judges and experts from across the world including the Committee to Protect Journalists, Reporters without Borders, the UN Special Rapporteur on Freedom of Expression and the International Bar Association's Human Rights Institute.

The International Bar Association's Human Rights Institute, serves as Secretariat to the High Level Panel of Legal Experts on Media Freedom, providing it with operational, technical, and legal assistance. The work of the High Level Panel, including the research contained in this report, has been supported by the Global Media Defence Fund, administered by UNESCO. At the time of publication, the Chair of the High Level Panel was Lord Neuberger of Abbotsbury and its Deputy Chairs were Professor Can Yeginsu and Ms. Catherine Amirfar.

Cover image: www.stock.adobe.com

www.ibanet.org/IBAHRI

Copyright © 2025 International Bar Association

3

Hate Speech

Amal Clooney and Alice Gardoll

I. Introduction	4	1.3. Genocide Convention Article 3	29
II. State Practice	7	1.4. Mandatory restrictions at the regional level	31
1. Overview of Hate Speech Laws	7	1.5. Summary	33
1.1. Type of speech	7	2. Discretionary Restrictions on Hate Speech	34
1.2. Harm	9	2.1. Legality	36
1.3. Intent	10	2.2. Legitimacy	38
1.4. Exclusions, exceptions and defences	11	2.2.1. Human Rights Committee	38
1.5. Penalties	11	2.2.2. European Court of Human Rights	39
2. Application of Hate Speech Laws Around the World	12	2.2.3. Inter-American Court and Commission	41
2.1. Europe	12	2.2.4. African Charter	41
2.1.1. The Netherlands	12	2.3. Necessity	44
2.1.1.1. Wilders	12	2.3.1. Harm	45
2.1.1.2. Basebya	13	2.3.2. Intent	47
2.1.2. Austria	14	2.4. Exclusions, exceptions and defences	49
2.2. Asia Pacific	16	2.4.1. Truth	49
2.2.1. Indonesia	16	2.4.2. Opinion	49
2.2.2. Myanmar	16	2.4.3. Public interest	52
2.2.3. Pakistan	17	2.4.4. 'Responsible journalism'	53
2.3. North and South America	17	2.5. Penalties	55
2.3.1. Canada	17	IV. Key Divergences in International and Regional Guidance	58
2.3.2. United States	19	V. Approach of Private Companies to Online Hate Speech	59
2.4. Middle East and Africa	22	VI. Recommendations	62
2.4.1. Tunisia	22		
2.4.2. Saudi Arabia	23		
III. International Legal Standards	23		
1. Mandatory Restrictions on Hate Speech	24		
1.1. ICCPR Article 20	26		
1.2. CERD Article 4	28		

I. Introduction

States must protect individuals from violence and discrimination, including when such harm is triggered by hateful speech.¹ And this is increasingly urgent: hate speech spreads faster and wider than ever before; hate-motivated crimes have risen to their highest levels in more than a decade in the United States, and an upsurge in hate crimes

¹ The authors and editors would like to thank Professor Evelyn Aswad, Professor Sarah Cleveland and Professor David Kaye for their expert advice on an earlier version of this chapter and for their invaluable respective contributions to this field.

has been recorded across the world.² But the right to speak freely is an essential right, the foundation of a free and democratic society and a prerequisite ‘for the full enjoyment of’ our other human rights.³ How should these rights be balanced?

International law provides an answer to this question in three human rights treaties—the International Covenant on Civil and Political Rights, the Convention on the Elimination of All Forms of Racial Discrimination and the Convention on the Prevention and Punishment of the Crime of Genocide—as well as regional human rights treaties covering Europe, Africa and the Americas. But unlike other types of speech, states can be *required* rather than simply permitted to penalize hate speech. Mandatory requirements to punish such speech—at least when it rises to the level of ‘incitement’—are contained in article 20 of the ICCPR, article 4 of CERD, article 3 of the Genocide Convention, as well as in one regional treaty, article 13(5) of the American Convention on Human Rights.⁴ But the harms that the speech must potentially incite vary, with ‘violence,’ ‘discrimination,’ or ‘hostility’ prohibited in the ICCPR; ‘violence,’ ‘discrimination,’ ‘hatred’ or ‘contempt’ prohibited under CERD;⁵ and ‘lawless violence’ or ‘other similar action’ outlawed in the American Convention provision.

The targeted group also diverges, with CERD only applying to hatred based on racial, national or ethnic origin, as opposed to broader discriminatory grounds recognized in other treaties.⁶ Finally, the penalties mandated by each treaty provision differ, with CERD, the Genocide Convention and the American Convention⁷ requiring criminalization, and the ICCPR requiring civil sanctions only.⁸ Many states have ratified more than one of these treaties, creating a web of obligations that adds to the complexity of states’ compliance with these standards.

These treaties also regulate the circumstances under which states can, but are not required to, penalize hateful speech. Although these standards are phrased in similar terms in the treaties, there has been some divergence in practice in this area.⁹ While the Human Rights Committee and the African and Inter-American regional human rights bodies have adopted similarly high standards to govern the permissible restrictions

² See, e.g., S. Haynes, ‘This Isn’t just a Problem for North America.’ The Atlanta Shooting Highlights the Painful Reality of Rising Anti-Asian Violence Around the World’ (Time, 22 March 2021). See also M. Yang, ‘Hate crimes in US rise to highest level in 12 years, says FBI report’ (The Guardian, 31 August 2021).

³ HRC, General Comment No. 34 (2011), §§2–4.

⁴ See ss III.1. (Mandatory Restrictions on Hate Speech) and III.1.4. (Mandatory restrictions at the regional level).

⁵ Although ‘contempt’ does not appear in the text of the treaty, it has been interpreted by the CERD Committee, the body with jurisdiction to interpret and apply the treaty, as being encompassed by its provisions. See CERD Committee, General Recommendation No. 35 (2013), §§13, 25. See s. III.1.2. (CERD Article 4).

⁶ See s. III.1.2. (CERD Article 4). See CERD Committee, General Recommendation No. 35 (2013), §13. Cf. ACHPR Art. 13(5); ICCPR Art. 20(2). See also ICCPR Art. 26.

⁷ The Inter-American Court has rejected the use of criminal law in cases involving public interest speech but has not ruled on cases of hate speech specifically. See IACtHR, *Álvarez Ramos v. Venezuela* (Series C, no. 380), 30 August 2019, §§120–122; s. III.2.5. (Penalties).

⁸ See s. III.1.1. (ICCPR Article 20), s. III.1.2. (CERD Article 4), s. III.1.4. (Mandatory restrictions at the regional level). See also HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, §10.4 (‘article 20 [ICCPR] does not expressly require the imposition of criminal penalties’).

⁹ E. Aswad & D. Kaye, ‘Convergence & Conflict: Reflections on Global and Regional Human Rights Standards on Hate Speech’ (2022) 20(3) *Northwestern Journal of Human Rights* 165, 186–209.

to speech, including stringent harm and intent requirements, the European Court of Human Rights has provided less protection to speech that is hateful.¹⁰ This divergent approach can be attributed in part to the European Court's 'margin of appreciation' doctrine, which provides states with discretion to determine whether a 'pressing social need' makes a speech restriction 'necessary' to promote one of the objectives listed in article 10(2) of the European Convention on Human Rights.¹¹ And article 17 of the European Convention, which bars abusive claims, has been construed as a 'guiltoline provision' that treats speech that is 'apt to destroy the fundamental values of the Convention' as being outside the scope of protection by the Court.¹² In other words, if the Court makes a threshold assessment that the speech at issue violates article 17, it will refuse to even consider whether it should be protected under the balancing test set out in article 10.

Around the world, laws regulating the punishment of hate speech are often vague and fail to address key elements such as the harm caused by the speech, the intent of the speaker and acceptable penalties for the speech. And yet such laws are usually criminal in nature.¹³ This creates an unclear level of protection for targets of hate speech and makes such laws particularly susceptible to abuse by authoritarian regimes that use them to silence dissent.

This chapter sets out the minimum international standards that apply to states drafting, interpreting and applying hate speech laws. It also recommends which approach is preferable from a policy perspective where there has been an omission or international bodies have taken divergent views. It recommends that European Court jurisprudence should move closer to the international, more speech-protective approach to hate speech, particularly when criminal laws apply. It also proposes that hate speech should only be subject to criminal penalties where it 'incites the commission of criminal offences or acts of violence' or 'similar action',¹⁴ ensuring that in other circumstances offensive and reprehensible speech is addressed using alternative means.¹⁵ And the chapter concludes that a more unified approach would make the international human rights framework a more practical guide to both states and private companies operating in this area.

¹⁰ See s. III.2. (Discretionary Restrictions on Hate Speech).

¹¹ See, e.g., ECtHR, *Zana v. Turkey* (App. no. 18954/91), 25 November 1997, §51; ECtHR (GC), *Stoll v. Switzerland* (App. no. 69698/01), 10 December 2007, §105, citing ECtHR, *Éditions Plon v. France* (App. no. 58148/00), 18 May 2004, §44; HRC, General Comment No. 34 (2011), §36.

¹² Article 17 provides that nothing in the Convention 'may be interpreted as implying ... any right to engage in any activity or perform any act aimed at the destruction of any of the rights and freedoms set forth herein'. See ECtHR, 'Factsheet—Hate Speech' (September 2023).

¹³ See s. II.1.5. (Penalties).

¹⁴ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §§4, 7. See s. III.1.4. (Mandatory restrictions at the regional level).

¹⁵ The conclusion that certain offensive speech should not be penalized through criminal or other laws is not intended to discourage robust action by both governments and private companies—including social media platforms—to ensure that racist, misogynistic, homophobic or other deplorable language is addressed, nor does it in any way condone this type of speech. See ch. 1 (Introduction).

II. State Practice

Broadly speaking, hate speech¹⁶ refers to speech that ‘attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are’: ‘in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor.’¹⁷ But ‘hate speech’ and ‘hate’ are not defined terms under international law—indeed, no human rights treaty uses this term.¹⁸ Many national laws also fail to provide a definition of the speech that is covered by legislation in this area. This is despite the fact that approximately 95 per cent of hate speech laws allow for the imposition of criminal penalties.¹⁹

1. Overview of Hate Speech Laws

Hate speech laws around the globe protect speech to very different degrees, both on paper and in practice. At one end of the spectrum, there are criminal offences proscribing speech without requiring any harm to have been intended or to have been likely to occur. At the other end are laws that require specific intent to cause genocide or violence as a direct result of the speech and either proof that such harm occurred or an objective likelihood that such harm was likely and imminent. Many states fail to define ‘hate,’²⁰ or to address the required mental state of the speaker, the harmful consequences of the speech or the foreseeable causal link between the two, resulting in overbroad statutes that chill speech and are susceptible to abuse by authorities seeking to stifle debate or dissent.

1.1. Type of speech

Some laws punish expression of a particular viewpoint, most commonly racism, without requiring any showing of culpability, criminal intent or foreseeable harm, or are unclear as to whether such elements must be established.²¹ Some states also proscribe the ‘display’, ‘dissemination’, ‘glorification’ or ‘promotion’ of hateful ideas without requiring any *mens rea* or showing of resulting harm. Examples include France criminalizing the

¹⁶ The review of state laws and practice on hate speech contained in this chapter does not aim to be comprehensive but rather to identify significant examples and trends. Therefore, where the language of a statute is set out, this does not necessarily include an analysis of all case law interpreting that language in the jurisdiction under review.

¹⁷ UN Strategy and Plan of Action on Hate Speech (2019), 2.

¹⁸ See s. III. (International Legal Standards).

¹⁹ See N. Alkiviadou, J. Mchangama & R. Mendiratta, *Global Handbook on Hate Speech Laws* (The Future of Free Speech 2020) (this figure is based on a survey of hate speech laws across 115 nations). The laws discussed in this State Practice section are all criminal unless otherwise indicated.

²⁰ See, e.g., Kenyan Constitution Art. 33(2); The European Union adopts a tautological definition: ‘[h]atred should be understood as referring to hatred based on race, colour, religion, descent or national or ethnic origin’ European Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law Chapeau §9. See s. II.2.3.1. (Canada).

²¹ See, e.g., Botswanan Penal Code Chapter 08:01 Art. 92 (fine-only offence).

display of Nazi badges, emblems or uniforms, with strict exemptions,²² and Austria penalizing anyone who ‘glorifies or extols the objectives’ of the Nazi party or ‘its institutions or actions.’²³

Case Study: Blasphemy Laws

At least 95 states have criminal laws that punish blasphemy: expression that insults or criticizes ‘religious doctrines, deities, symbols’ or the ‘feelings’ of religious communities.²⁴ Although decriminalized in many Western countries, blasphemy remains an offence in around 40 per cent of the world’s states and carries the death penalty in at least five countries.²⁵

The crime of blasphemy is often expressed in vague terms that cast a wide net over speech that may insult religious ideas or groups. Broad blasphemy laws include the criminalization of ‘contempt of heavenly religions,’²⁶ speech that ‘provide[s] an opinion or disperse[s] information’ on a matter ‘which has conflicting opinions among Islamic Scholars’ and words that ‘preach or spread any religion other than Islam.’²⁷

Other states include religious grounds within a broader hate speech law. For instance, Brazil criminalizes ‘injuring someone, offending their dignity or decorum’ with additional penalties ‘[i]f the injury consists in the use of elements referring to race, color, ethnicity, *religion*, origin or condition of elderly or disabled person[s].’²⁸

Case Study: ‘Denial laws’

Denial laws prohibit speech denying the truth of historical facts related to genocide, crimes against humanity or other international crimes. Several European

²² French Penal Code R. 645-1 (except for purposes of a film, show or exhibition involving historical invocation, punishable by fine, and additional penalty including the ban to carry an arm, the confiscation of one’s arm and community service).

²³ Austrian National Socialism Prohibition Act Art. 3d (punishable by up to 10 years’ imprisonment, or if ‘the activity should pose a particularly grave danger’, up to 20 years’ imprisonment).

²⁴ See United States Commission on International Religious Freedom, ‘Legislation Factsheet: Blasphemy’ (September 2023).

²⁵ Ibid. (citing Mauritania, Brunei, Iran and Pakistan). Saudi Arabia does not have a written penal code and relies on judges’ interpretation of Sharia law, with judges previously interpreting apostasy as a capital offence: see, e.g., Human Rights Watch, ‘Saudi Arabia: Poet Sentenced to Death for Apostasy’ (23 November 2015). See IBAHRI, *On Religious Freedom and Discontent: Report on International Standards Relating to Blasphemy Laws and Media Freedom* (2023). In Pakistan, s. 295-C of the Penal Code 1860 was introduced in 1986 which criminalizes blasphemy against the Holy Prophet, punishable by death or life imprisonment. In 1990, the Federal Shariat Court declared that the only penalty under s. 295-C would be the death penalty. Although no executions have taken place under the law, suspects are often killed by vigilantes on the way to court. See s. II.2.2.3. (Pakistan). See also Pew Research Center, ‘Four-in-ten countries and territories worldwide had blasphemy laws in 2019’ (25 January 2022).

²⁶ Egyptian Criminal Code s. 98(f) (punishable by up to five years’ imprisonment).

²⁷ Maldivian Regulation on Protecting Religious Unity of Maldives Citizens (2011/R-40) ss 5, 6.

²⁸ Brazilian Decree-Law No. 2.848 of 1940 Penal Code Art. 140 (punishable by up to three years’ imprisonment).

states as well as Israel proscribe expression denying the Holocaust,²⁹ and certain denial laws are also combined with broader ‘memory laws’ that compel adherence to a particular historical narrative.³⁰ This is the case, for instance, in Russia, where it is an offence to ‘deny facts recognized by the international military tribunal that judged and punished the major war criminals of the European Axis countries, to approve of the crimes this tribunal judged, and to spread intentionally false information about the Soviet Union’s activities during World War II.’³¹ An EU directive has sought to harmonize such laws by requiring that EU states punish the ‘public condoning, denying, and grossly trivializing’ crimes described in the Rome Statute of the International Criminal Court (genocide, war crimes and crimes against humanity) and the Nuremberg Charter (Holocaust denial) wherever they occur.³²

These laws are usually content-based, criminalizing ‘denialism’ without further requirements to establish any harmful consequences.³³ Some states do require additional elements, however. This includes Croatia, which punishes public approval, denial or gross trivialization of genocide, crimes against humanity, war crimes, or crimes of aggression when ‘directed ... in a manner likely to incite to violence or hatred against such a group or a member of such a group,’³⁴ and Germany, which proscribes the denial of specific crimes ‘in a manner suitable for causing a disturbance of the public peace.’³⁵

1.2. Harm

States’ hate speech laws differ based on the nature of the harm required, if any, and the likelihood or imminence of any harm that may result from the speech. At one end of the spectrum, some states’ laws require that speech directly incites genocide in order to be criminalized. Other laws recognize a broader range of harms or omit a harm requirement altogether. For instance, some nations punish speech that incites any ‘violence,’³⁶

²⁹ See, e.g., French Law of 13 July 1990 providing for the punishment of all racist, anti-Semitic, and xenophobic acts, also known as the Gayssot Act and amending French Law of 29 July 1881 on the Freedom of the Press with a new Art. 24bis (punishing the denial of the Holocaust by up to one year imprisonment and/or a fine of 45,000 Euros). See s. III.2.4.2. (Opinion); Belgian Act of 23 March 1995 Art. 1 (concerning the denial, gross minimization, justification or approval of the genocide, in the sense of the Geneva Convention of 9 December 1948, by the German National Socialist Regime during the Second World War (punishable by up to one year imprisonment and a fine up to 5,000 EUR)); see generally M. Whine, ‘Expanding Holocaust Denial and Legislation Against It’ (2008) 20(1–2) *Jewish Political Studies Review* 1. Israeli Denial of Holocaust (Prohibition) Law 5746-1986 ss 2, 4 (denial must be accompanied by ‘intent to defend the perpetrators of those acts or to express sympathy or identification with them,’ punishable by up to five years’ imprisonment).

³⁰ See, e.g., four different memory laws currently in force in France: (i) the 1990 Gayssot Act, (ii) the Law No. 2001-70 of 29 January 2001 relating to the recognition of the Armenian Genocide of 1915, (iii) the Law No. 2001-434 of 21 May 2001 recognizing the slave trade and slavery as a crime against humanity (also known as the Taubira Act) and (iv) Law No. 2005-158 of 23 February 2005 on the recognition of the Nation and national contribution in favour of repatriated French people.

³¹ See Russian Penal Code Art. 354.1 (punishable by up to three years’ imprisonment).

³² European Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law Art. 1. See also Art. 1(4).

³³ However, for at least some of these laws, the harmful consequence is presumed to result per se from the content of the speech. See HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996 (where the HRC upheld France’s rationale for the Gayssot Act). Cf. HRC, General Comment No. 34 (2011), §49 (which appears to overrule *Faurisson*): see s. III.2.4.2 (Opinion).

³⁴ Croatian Penal Code Art. 325(4) (punishable by up to three years’ imprisonment).

³⁵ German Criminal Code Art. 130 (punishable by up to three years’ imprisonment).

³⁶ See, e.g., Malta Criminal Code Chapter 9, Art. 82A (punishable by up to 18 months’ imprisonment).

or promotes ‘hostility.’³⁷ And Canada provides criminal liability for a defendant who ‘advocates or promotes genocide’—without a separate showing of harm.³⁸

Some states have a stringent requirement relating to both the likelihood that harm will occur as the result of ‘hate speech’ and how imminent that harm is likely to be. For instance, US law only allows criminalization of speech that is ‘directed to inciting or producing imminent lawless action and is likely to incite or produce such action.’³⁹ Latvian law requires that substantial harm *has been caused*.⁴⁰ And Estonian law criminalizes incitement to ‘hatred, violence or discrimination on the basis of nationality, race, colour, sex, language, origin, religion, sexual orientation, political opinion, or financial or social status’ only ‘*if this results in danger to the life, health or property of a person*.’⁴¹ But many states’ laws are vague as to both likelihood and imminence, place the bar very low or fail to address these issues entirely.⁴²

1.3. Intent

The EU regulation on hate speech states that the legal prohibition of hate speech should ‘ensure that . . . *intentional* conduct is punishable.’⁴³ But it is not clear what form of intent is required and significant variation exists in hate speech laws, even within Europe. For instance, Ireland’s law penalizes words that are ‘threatening, abusive or insulting *and are intended or, having regard to all the circumstances, are likely to stir up hatred*.’⁴⁴ The United Kingdom,⁴⁵ Kenya and Malta also have laws that refer to intent.⁴⁶ Other states, like the Netherlands, use a knowledge test, requiring that the person makes public a

³⁷ Malawian Preservation of Public Security Act Chapter 14:02 Arts 4(c), 14; Ugandan Penal Code Act 1950 s. 41(1); Act of 10 May 2007 aimed at combating certain forms of discriminations Art. 20; French Law on the Freedom of the Press of 29 July 1881 Art. 24(5).

³⁸ Canadian Criminal Code s. 318. See also Australian Criminal Code Act 1995 s. 80.2D (punishable by up to seven years’ imprisonment).

³⁹ US Supreme Court, *Brandenburg v. Ohio* 395 U.S. 444, 9 June 1969, 477. See also other limited cases in which speech may be restricted: See s. II.2.3.2. (United States).

⁴⁰ Latvian Criminal Law s. 150 (punishable by up to four years’ imprisonment in certain circumstances).

⁴¹ Estonian Penal Code RT I 2001, 61, 364 s. 151 (emphasis added) (punishable by a fine or detention of up to three years if the act causes death, damages to health or serious consequences, or is committed by someone previously punished for such an act).

⁴² See, e.g., Greek Law No. 4285/2014, amending Law No. 927/1979, Art. 1.1 (speech that ‘*may lead to violence, hatred [or] discrimination . . . in a way that may endanger public order*’) (punishable by up to three years’ imprisonment) (emphasis added). See also Greek Law No. Law 4491/2017.

⁴³ European Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law, Art. 1(1) (emphasis added).

⁴⁴ Irish Prohibition of Incitement to Hatred Act 1989, ss 2, 3, 6(b) (punishable by up to two years’ imprisonment and a fine of 11,650 EUR when charged on indictment) (emphasis added). The proposed Criminal Justice (Incitement to Violence or Hatred and Hate Offences) Bill 2022 would—if it is enacted—repeal the Irish Prohibition of Incitement to Hatred Act 1989, and introduce a new hate crime offence. The proposed offence introduces a recklessness standard and increases the maximum custodial sentence on indictment from two to five years. See Coalition Against Hate Crime, ‘Coalition Against Hate Crime Second Stage Briefing Note for Senators on the Criminal Justice (Incitement to Violence or Hatred and Hate Offences) Bill 2022’ (8 June 2023).

⁴⁵ Unless otherwise indicated, references to the United Kingdom and the UK refer to England and Wales, Scotland and Northern Ireland. Although many of the same laws apply across these jurisdictions, these laws can vary and Scotland has a separate legal system.

⁴⁶ UK Public Order Act 1986 s. 18(1) (punishable by up to seven years’ imprisonment when charged on indictment). See also UK Public Order Act 1986, ss 19, 20, 21, 22. The Hate Crime and Public Order (Scotland) Act 2021 will repeal ss 18–22 of the UK Public Order Act 1986 with respect to Scotland—once the relevant provisions are brought into force—and create a broader hatred offence which also refers to intent. See also Kenyan National Cohesion and Integration Act 2008 s. 13 (punishable by up to three years’ imprisonment), Maltese Criminal Code Chapter 9 Art. 82A.

statement ‘which he knows or should reasonably suspect to be insulting to a group of persons because of their race’, or omit a culpability requirement entirely.⁴⁷

1.4. Exclusions, exceptions and defences

Some states have a number of alternative defences available to hate speech crimes; others have one or none. Canada provides defences to the charge of wilful promotion of hatred where a defendant establishes that the statements were (a) ‘true’, related to ‘a religious subject or text’ and disseminated in ‘good faith’; or (b) were in the ‘public interest’, for the ‘public benefit’ and the speaker had ‘reasonable grounds to believe the truthfulness’ of the statement.⁴⁸ The United Kingdom has a limited good faith defence, whereby the defendant will not be guilty of an offence ‘if he did not intend his words . . . to be, and was not aware that [they] might be, threatening, abusive or insulting’.⁴⁹ Other laws do not apply to academic debate or news reports that are published to criticize the underlying hatred.⁵⁰

1.5. Penalties

A survey of 115 countries’ hate speech laws revealed that over 95 per cent of states use criminal—instead of or in addition to civil—penalties for hate speech.⁵¹ The range of criminal penalties, however, varies considerably across jurisdictions. For instance, although some laws have hate speech offences that result only in fines or carry very short custodial sentences,⁵² many others provide for substantial sentences of up to five years’ imprisonment,⁵³ or even lengthier sentences,⁵⁴ and at least five authorize capital punishment for blasphemous speech.⁵⁵

Some penalties for hate speech are specific to media companies and online content. In nations such as Bolivia, publication of discriminatory ideas can also result in the suspension of a media company’s operating licence.⁵⁶ Elsewhere, states impose liability or

⁴⁷ See, e.g., Dutch Criminal Code s. 137(e) (punishable by up to one year imprisonment). Dutch law also criminalizes intentional hate speech: see Dutch Criminal Code s. 137(c).

⁴⁸ A final defence from prosecution arises if ‘in good faith’, the speaker referred to hate speech ‘for the purpose of removal’ of such speech: Canadian Criminal Code s. 319(3).

⁴⁹ UK Public Order Act 1986 s. 18(5). This defence only applies if the person ‘is not shown to have intended to stir up racial hatred’. See also ss 19(2), 20(2), 21(3), 22(5) and (6); Irish Prohibition of Incitement to Hatred Act 1989 s. 2(a).

⁵⁰ See, e.g., Montenegrin Law on Media, Official Gazette of Montenegro 82/2020 dated 6 August 2020; Dutch Criminal Code s. 137(e) (excluding culpability when the speech is made ‘for any reason other than the provision of factual information’).

⁵¹ See Alkiviadou, Mchangama, & Mendiratta (n 19).

⁵² See, e.g., Rwandan Law Determining Offences and Penalties in General No. 68 2018 Art. 154 (punishable by three months’ imprisonment).

⁵³ See, e.g., Canadian Criminal Code Art. 318 (five years’ imprisonment for advocating or promoting genocide) and Art. 319 (two years’ imprisonment for hate speech); Brazilian Law 9459 Art. 20 (five years’ imprisonment when committed through media or publishing).

⁵⁴ Austrian National Socialism Prohibition Act Art. 3d. Austria has since passed other hate speech provisions with less stringent penalties, such as Art. 188 of the Criminal Code, which provides for up to six months’ imprisonment for insulting religion.

⁵⁵ Mauritania, Brunei, Iran, Pakistan and Saudi Arabia (see United States Commission on International Religious Freedom, ‘Legislation Factsheet: Blasphemy’ (September 2023) and Human Rights Watch, ‘Saudi Arabia: Poet Sentenced to Death for Apostasy’ (23 November 2015)). See s. II.1.1. (Type of speech) (Case Study: Blasphemy Laws).

⁵⁶ Bolivian Law Against Racism and All Forms of Discrimination 2010 Art. 16 (punishable by suspension of operating licence or financial penalties). See also German Civil Code ss 823(2), 826 (providing civil remedies where a

hefty penalties for internet intermediaries when hate speech is found on their platform. A controversial example is Germany's 2017 Network Enforcement Act (commonly known as NetzDG) which compels social media platforms to remove unlawful content, including hate speech, from accounts that have more than two million users.⁵⁷ Under this law, removal must take place for 'manifestly unlawful' content within 24 hours, and for all other unlawful content within seven days, failing which penalties of up to 50 million Euros can be imposed.⁵⁸ UN and media freedom groups have expressed concern about the breadth and severity of this law, noting that companies have 'little incentive to err on the side of free expression.'⁵⁹ And more than a dozen countries, including Russia and Turkey, have since passed similar laws.⁶⁰

2. Application of Hate Speech Laws Around the World

Trends in the implementation of hate speech laws vary across geographic regions. Many European hate speech cases have involved balancing a desire to prevent the rise of racist right-wing sentiment with the need to protect freedom of expression and political debate. In the Asia Pacific and Middle Eastern context, certain repressive governments have used vague hate speech provisions and blasphemy laws as a pretext to chill political dissent. American jurisprudence has traditionally reflected a high watermark of free speech in this space, where courts will accept the regulation of hateful speech only in very limited circumstances, namely where such speech is likely to incite imminent violence or unlawful action.⁶¹

2.1. Europe

2.1.1. *The Netherlands*

2.1.1.1. Wilders A series of high-profile cases against Geert Wilders, leader of the Dutch far-right 'Party for Freedom', illustrates where Dutch courts draw the line in balancing the right to speak and the right not to be the target of racist and xenophobic sentiment.⁶²

In the first case, criminal charges were brought for statements Wilders made during media interviews and on his party website related to Muslim immigrants and

person 'commits a breach of a statute that is intended to protect another person' or 'intentionally inflicts damage on another person', including incitement offences).

⁵⁷ German Network Enforcement Act 2017 ss 1, 3. This entered into force on 1 October 2017.

⁵⁸ *Ibid.*, s. 3; Human Rights Watch, 'Germany: Flawed Social Media Law' (14 February 2018).

⁵⁹ Human Rights Watch, 'Germany: Flawed Social Media Law' (14 February 2018); UN Special Rapporteur D. Kaye, Promotion and protection of the right to freedom of opinion and expression (2019) UN Doc. A/74/486, §32.

⁶⁰ 'Censorious governments are abusing "fake news" laws' (The Economist, 13 February 2021). Cf. EU Digital Services Act (directly applicable from early 2024, including fines up to six per cent of the global turnover of a service provider).

⁶¹ See s. II.2.3.2. (United States).

⁶² See s. III.1.1. (ICCPR Article 20).

Moroccan residents in the Netherlands. In 2011, the Amsterdam District Court acquitted Wilders of five charges including insult and ‘incitement to hatred and discrimination’ on grounds of religion and race.⁶³ The Court considered statements like ‘[w]e have to stop the tsunami of Islamization’ to be criticisms of Islam as a religion, rather than a particular group of people, and held that the ‘mere circumstance that offensive statements about a religion also offend the adherents of that religion is . . . not sufficient’ to make out the offence.⁶⁴

In contrast, in a second case, Wilders was convicted for asking an audience at a televised election rally: ‘do you want more or less Moroccans in this city and the Netherlands?’ and, in compliance with earlier instructions, the audience chanted back ‘less, less, less,’ following which Wilders replied ‘well, then we’ll arrange that.’ The Dutch courts⁶⁵ held that he could be convicted for ‘insult’ but that a prison term or fine was not an appropriate punishment.⁶⁶ It also held that he could not be convicted of incitement to racial discrimination because there was no evidence that he intended to incite others to hate or discriminate against Moroccans or that he ‘knowingly accepted the significant chance that the public . . . would be moved to do so.’⁶⁷ The court considered that Wilders was instead seeking ‘political gain’ from his statement and as a result should be acquitted of this charge.⁶⁸

2.1.1.2. Basebya In 2013, Yvonne Basebya became the first Dutch citizen to be convicted of incitement to genocide and was sentenced to nearly seven years’ imprisonment.⁶⁹ At trial, the prosecution presented evidence that Basebya was a member of

⁶³ Dutch Court of Amsterdam, *Judgment of the District Court of Amsterdam in the criminal case against Geert Wilders* Case Number 13-425046-09, 23 June 2011.

⁶⁴ Dutch Court of Amsterdam, *Judgment of the District Court of Amsterdam in the criminal case against Geert Wilders* Case Number 13-425046-09, 23 June 2011, §4.2; de Rechtspraak, ‘Verdict of the Amsterdam district court as regards the Wilders trial’ (23 June 2011). The Court also held that incitement to hatred and discrimination on grounds of ‘race’ could not apply to speech directed to Moroccans or non-Western immigrants. Dutch Court of Amsterdam, *Judgment of the District Court of Amsterdam in the criminal case against Geert Wilders* Case Number 13-425046-09, 23 June 2011, §4.4. The Dutch Public Prosecution Service had refused to prosecute Wilders because it did not consider his statements illegal, but was later ordered by the Court of Appeal upon a complaint to pursue the prosecution. During the process, the Public Prosecution Service argued that Wilders should be acquitted on all counts and did not appeal the acquittal.

⁶⁵ The first instance court found Wilders guilty on two counts of insult and inciting racial discrimination: Dutch District Court of the Hague, *State of the Netherlands v. Wilders* Case Number 09/837304-15, 9 December 2016, §5.4.3. The Court came to a different decision on the legal definition of race than in Wilders’ earlier proceedings, interpreting the term consistently with the CERD to encompass persons linked by their ‘common Moroccan descent’. *Ibid.*, §5.4.2. Wilders was acquitted of two separate counts which related to statements made on another day saying that people should hope for ‘if possible, less Moroccans’ in Dutch cities, as the Court found that these statements were made spontaneously at a market and he ‘did not have the (conditional) intent to insult people of Moroccan descent or to incite discrimination or hatred’: §5.4.4.

⁶⁶ Dutch Court of Appeal in The Hague, *Wilders v. State of the Netherlands* Case Number 22-000007-17, 4 September 2020, §8.4. The Appeals Court agreed with the lower Court that a conviction was sufficient punishment (even though a maximum two-year prison term was available) and this judgment was upheld by the Supreme Court. *Ibid.*, §13; Dutch Supreme Court: *Wilders v. State of the Netherlands* Case Number 20/03005, 6 July 2021.

⁶⁷ Dutch Court of Appeal in The Hague, *Wilders v. State of the Netherlands* Case Number 22-000007-17, 4 September 2020, §8.5.

⁶⁸ *Ibid.* The District Court had already ruled that incitement to hatred required an additional reinforcing component, beyond merely strong rhetoric, which instigates people to take action that Wilders lacked: Dutch District Court of The Hague, *State of the Netherlands v. Wilders* Case Number 09/837304-15, 9 December 2016, §5.4.3.2.

⁶⁹ ‘Rwandan-born Dutch woman jailed for inciting genocide’ (The Guardian, 1 March 2013); Dutch District Court of The Hague, *The Prosecutor v. Yvonne Basebya* Case Number 09/748004-09, 1 March 2013.

an extremist pro-Hutu political party, and a wealthy and well-respected member of her community. She was alleged to have acted as an ‘animator’ at meetings where she ‘aggressively instilled hatred against the Tutsis and ... encouraged members to kill Tutsis’, for example by leading the singing of an extremist anti-Tutsi song entitled ‘Tubatsembatsembe’, meaning ‘let’s exterminate them all.’⁷⁰ The court heard evidence that following these meetings, Tutsis felt it was dangerous to be on the streets in Basebya’s neighbourhood in Kigali, and that they would remain at home.⁷¹

The District Court of The Hague held that for this conduct to amount to incitement to genocide, the prosecution must prove that there was intent to ‘entirely or partially exterminate the Tutsi-population group’; and that the impugned speech was a ‘direct’ call for genocide.⁷² The District Court considered that ‘such intent must be ‘firmly established’ with ‘no other explanation ... possible.’⁷³ In light of the circumstances, namely the ‘political framework in which the accused’s remarks were made’, the violent attacks against Tutsis at the time, Basebya’s social standing in the community and the repeated nature of her conduct, the District Court considered that ‘no other conclusion’ was possible than that Basebya ‘had the aim to destroy the Tutsi-population group’, by leading the singing of the song at the meetings.⁷⁴ In addition, the element of directness was held to be satisfied on the basis that ‘[i]n view of the nature of the meetings and the public that attended them, this form of expression could not be explained in any other way than that it was an explicit call for the extermination of the local Tutsi-population.’⁷⁵ Basebya was sentenced to the maximum sentence of six years and eight months’ imprisonment for incitement to genocide.⁷⁶

2.1.2. Austria

In 2013, Austria’s Supreme Court confirmed the conviction of a woman known as ‘E.S.’, for disparaging religious doctrines on the basis of her description of Prophet Muhammad as a paedophile during a seminar.⁷⁷

E.S. had held a number of seminars at the Freedom Party Education Institute, open to members of the Freedom Party, guests and the public. During some of these seminars, attended by an undercover journalist, E.S. stated that ‘one of the biggest problems

⁷⁰ Dutch District Court of The Hague, *The Prosecutor v. Yvonne Basebya* Case Number 09/748004-09, 1 March 2013, §12(8) and (10).

⁷¹ *Ibid.*, §12(16).

⁷² *Ibid.*, §12(8). The Court also held that two other factors must be proved: (1) that the Tutsi population was a protected group and (2) that the speech was made publicly. The Court relied on the Rwandan Tribunal’s finding that the Tutsi population group can be regarded as an ethnic group within the meaning of the Genocide Convention, citing ICTR, *Prosecutor v. Kayishema and Ruzindana* (ICTR-95-1-T), Trial Judgement, 21 May 1999, §§522–526.

⁷³ Dutch District Court of The Hague, *The Prosecutor v. Yvonne Basebya* Case Number 09/748004-09, 1 March 2013, §12(10), (12).

⁷⁴ *Ibid.*, §12(13)–(20). See s. III.1.3. (Genocide Convention Article 3).

⁷⁵ *Ibid.*, §12(24). The District Court determined that the speech was ‘public’, as the meetings did not have a ‘closed character’, could be heard from inside witnesses’ homes and seen from the public road. *Ibid.*, §12(29). She was acquitted of the remaining charges including committing genocide.

⁷⁶ Both the prosecutor as well as the defendant initially appealed the decision, but later withdrew their appeals which made the District Court’s judgment final. In May 2015 the Supreme Court rejected Basebya’s request for revision; see Dutch Supreme Court, *Revision* Case Number 14/05251, 13 May 2015.

⁷⁷ Austrian Supreme Court, *Decision on request for renewal* Reference Number 15Os52/12d, 11 December 2013.

we are facing today is that Muhammad is seen as the ideal man, the perfect human, the perfect Muslim ... [but] he was a warlord, he had many women ... and liked to do it with children'. She also stated: '[a] 56-year-old and a six-year-old? ... What do we call it, if it is not paedophilia?'⁷⁸ These comments referred to the Prophet Muhammad's marriage to a child named Aisha who, according to religious tenets, was six at the time of the marriage and nine at the time it was consummated.⁷⁹

E.S. was convicted of an offence on the basis that her speech 'publicly disparages or insults a person who ... is an object of veneration of a ... religious community' in circumstances where their 'behaviour is likely to arouse justified indignation', a crime punishable by up to six months' imprisonment in Austria.⁸⁰ E.S. argued that her speech 'merely criticized the notion that an adult had had sexual intercourse with a nine-year-old and questioned whether this amounted to paedophilia' and should be protected.⁸¹ However, three levels of Austrian courts—including the Supreme Court of Austria—disagreed, finding that such speech was not protected by article 10 of the European Convention on Human Rights. E.S. was ordered to pay costs and a fine of 480 Euros, with 60 days' imprisonment in default.

Austria's Supreme Court held that this law had a legitimate aim that was 'the protection of religious peace and the religious feelings of others', and that a criminal conviction may be necessary to pursue this aim.⁸² The court held that the speech was likely to arouse 'justified indignation' as it was neither 'a factual criticism of religion' nor a contribution to the public debate on questions of child marriage or Islam, but rather was 'insulting' and made 'deliberately degrading remarks about the Prophet Mohammed'.⁸³ Although the Supreme Court accepted the lower court's finding that 'criticising child marriages was justifiable', it highlighted that E.S. had improperly disregarded the fact that 'the marriage had continued until the Prophet's death, when Aisha had already turned eighteen'.⁸⁴ The Supreme Court also considered a criminal conviction to be 'proportionate'.⁸⁵ This decision, and its endorsement by the European Court of Human Rights, has been the subject of extensive criticism on the basis that it was poorly reasoned and insufficiently protective of speech.⁸⁶

⁷⁸ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §13.

⁷⁹ *Ibid.*, §14.

⁸⁰ Austrian Criminal Code Art. 188. Charges were originally brought against E.S. for inciting hatred, pursuant to Art. 283(1) of Austria's Criminal Code (punishable by fine or imprisonment of up to two years).

⁸¹ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §16.

⁸² Austrian Supreme Court, *Decision on request for renewal* Reference Number 15Os52/12d, 11 December 2013.

⁸³ The Court of Appeal also held that there was a distinction between a child marriage and paedophilia and that there was no reliable evidence to suggest that Muhammad's 'primary sexual interest in Aisha had been her not yet having reached puberty'. ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §18. See also Austrian Supreme Court, *Decision on request for renewal* Reference Number 15Os52/12d, 11 December 2013, 4, 13, 14.

⁸⁴ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §§14, 22; Austrian Supreme Court, *Decision on request for renewal* Reference Number 15Os52/12d, 11 December 2013.

⁸⁵ Austrian Supreme Court, *Decision on request for renewal* Reference Number 15Os52/12d, 11 December 2013, cited in ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §14.

⁸⁶ See s. III.2.2. (Legitimacy) (Case Study: International Standards on Blasphemy Laws) See, e.g., M. Milanovic, 'Legitimizing Blasphemy Laws Through the Backdoor: The European Court's Judgment in *E.S. v. Austria*' (EJIL: Talk!, 29 October 2018); C. Yeginsu & J. Williams, 'Criminalizing Speech to Protect Religious Peace? The ECtHR Ruling in *E.S. v. Austria*' (Just Security, 28 November 2018).

2.2. Asia Pacific

2.2.1. Indonesia

In Indonesia, a number of prosecutions demonstrate how hate speech laws without robust harm and intent requirements can be used to stifle expression. One such law criminalizes speech ‘aimed at inflicting hatred or dissension on individuals and/or certain groups’ based on ethnicity, race or religion.⁸⁷ In 2012, civil servant Alexander Aan was convicted under the law and sentenced to two and a half years’ imprisonment for posts in an atheist Facebook group that explained why he did not believe in God’s existence. He was released on parole after serving 19 months in prison.⁸⁸

More recently, Ahmad Dhani, an Indonesian musician and political figure, was sentenced to one year of imprisonment under the same provision.⁸⁹ The speech at issue was a series of tweets regarding the former Jakarta governor, then on trial, including a tweet stating that ‘anyone who supports’ the governor ‘is scum and deserves to be spat [on] in the face.’⁹⁰ The decision has been criticized on the basis that ‘none of the tweets in question mention the governor’s ethnicity, religion or race—as specifically required’ in the law and that the ‘elasticity of these laws advantage whoever is in power.’⁹¹ Although Indonesia’s President has asked police officers to be ‘selective’ when using this law, it remains on the books in Indonesia.⁹²

2.2.2. Myanmar

Myanmar has used hate speech laws to stifle expression on a very broad basis. This includes the conviction of three employees and managers of the ‘V Gastro Bar’ in Yangon for posting a Facebook advertisement that showed a Buddha wearing headphones, with accompanying text advertising ‘bottomless frozen’ cocktails.⁹³ Although the post was deleted and an apology was issued, it was reshared on social media and prompted an outcry from Buddhist nationalist groups. The bar owner and managers were subsequently charged with insulting ‘religion’ or ‘religious beliefs’ with ‘deliberate and malicious intention.’⁹⁴

The men were convicted and sentenced to over two years’ imprisonment and hard labour.⁹⁵ The court reasoned that ‘it is clear’ that the speech ‘offended the majority religion in this country’ and that ‘ignorance of the law is not an excuse’, holding that the bar

⁸⁷ Indonesian Law Concerning Electronic Information and Transactions Arts 27(3), 28(2).

⁸⁸ J. Cochrane, ‘Embrace of Atheism Put an Indonesian in Prison’ (The New York Times, 3 May 2014).

⁸⁹ Indonesian Law Concerning Electronic Information and Transactions Arts 27(3), 28(2).

⁹⁰ ‘Ahmad Dhani walks free after serving 11 months for hate speech’ (The Jakarta Post, 30 December 2019).

⁹¹ See T. Paterson, ‘Indonesian Cyberspace Expansion: A Double-edged Sword’ (2019) 4 *Journal of Cyber Policy* 216.

⁹² ‘Jokowi asks the police to use the ITE law wisely’ (Netral.News, 16 February 2021); Office of Assistant to Deputy Cabinet Secretary of the Republic of Indonesia, ‘President Jokowi: ITE Law Must Fulfill Public Sense of Justice’ (15 February 2021).

⁹³ See, e.g., Columbia University Global Freedom of Expression Database, ‘The Case of V Gastro Bar (Philip Blackwood, Htut Ko Ko Lwin and Tun Thurein)’; W. Moe & A. Ramzy, ‘Myanmar Sentences 3 to Prison for Depicting Buddha Wearing Headphones’ (The New York Times, 17 March 2015).

⁹⁴ Columbia University Global Freedom of Expression Database, ‘The Case of V Gastro Bar’ (n 93).

⁹⁵ See Myanmar Penal Code Art. 295. An additional charge related to disobeying a civil servant under Art. 188 and their sentences reflected the conviction on both charges. A charge of destroying, damaging or defiling a place of worship or sacred object was dropped.

manager—a man from New Zealand—should have known better since he had ‘stayed in Myanmar for more than three years.’ And the bar’s local owner ‘failed to instruct his foreign staff about the culture and traditions of the country.’⁹⁶ The defendants’ lawyer declined to comment on the verdict, stating ‘if I make any comment on the court’s decision I will end up in jail.’⁹⁷ The verdict was sharply criticized⁹⁸ and the bar manager was released following a presidential pardon after serving almost a year in prison, with his co-defendants reportedly released at the same time.⁹⁹

2.2.3. *Pakistan*

In August 2021, an eight-year-old boy was charged under Pakistan’s blasphemy laws, becoming the youngest person ever charged with blasphemy in that country, an offence which carries the death penalty.¹⁰⁰ He was accused of urinating in the library of a madrassa, and his release on bail prompted an attack on a Hindu temple. Police dropped the charges against him within days after media and government pressure.¹⁰¹ Although no executions have taken place under the law since the death penalty was introduced for the crime in 1986, a number of defendants remain on death row or in solitary confinement,¹⁰² and suspects are often attacked and in some instances killed by mobs.¹⁰³ Reports suggest at least 65 people have been killed over claims of blasphemy since 1990, and that blasphemy charges are often brought against religious minorities as a mechanism to settle personal disputes.¹⁰⁴ And in January 2023, lawmakers expanded Pakistan’s blasphemy laws to include insult to the Prophet’s wives, companions or close relatives, punishable by up to life imprisonment.¹⁰⁵ Blasphemy charges were also made a non-bailable offence.¹⁰⁶

2.3. North and South America

2.3.1. *Canada*

The landmark decision of the Supreme Court of Canada in *R. v. Keegstra* upheld Canada’s hate speech provisions as consistent with the free speech provisions enshrined in Canada’s Charter of Rights and Freedoms.¹⁰⁷ Keegstra was a high school teacher who

⁹⁶ Columbia University Global Freedom of Expression Database, ‘The Case of V Gastro Bar’ (n 93).

⁹⁷ J. Callil, ‘A Prison Sentence for a Facebook Image Shows How Restrictive Burma’s Anti-Free-Speech Laws Have Become’ (Vice, 24 March 2015) citing the Myanmar Times.

⁹⁸ Amnesty International, ‘Myanmar: “Buddha bar” guilty verdict another blow to freedom of expression’ (17 March 2015).

⁹⁹ ‘Phil Blackwood en route to New Zealand after release’ (Frontier Myanmar, 28 January 2016).

¹⁰⁰ H. Janjua, ‘Eight-year-old becomes youngest person charged with blasphemy in Pakistan’ (The Guardian, 9 August 2021).

¹⁰¹ H. Janjua, ‘Pakistan police drop blasphemy charges against eight-year-old’ (The Guardian, 12 August 2021).

¹⁰² Human Rights Watch, ‘Pakistan: End Ordeal for ‘Blasphemy’ Defendants’ (6 October 2019). Defendants continue to receive death penalty sentences: see, e.g., A. Hashim, ‘Pakistan court sentences woman to death for WhatsApp “blasphemy”’ (Al Jazeera, 20 January 2022).

¹⁰³ United States Commission on International Religious Freedom, *Violating Rights: Enforcing the World’s Blasphemy Laws* (2020).

¹⁰⁴ Human Rights Watch, ‘Pakistan: End Ordeal for ‘Blasphemy’ Defendants’ (6 October 2019).

¹⁰⁵ S. Masood, ‘Pakistan Strengthens Already Harsh Laws Against Blasphemy’ (The New York Times, 21 January 2023).

¹⁰⁶ Ibid.

¹⁰⁷ Canadian Supreme Court, *R v. Keegstra* [1990] 3 S.C.R. 697, 13 December 1990 is the leading case in a trilogy of landmark Supreme Court decisions affirming the constitutional validity of hate speech limitations, including *R*

expressed anti-Semitic statements to his students, including describing Jewish people as ‘subversive’ and suggesting that they ‘created the Holocaust to gain sympathy’. He requested that students reproduce his teachings and would mark them down in class if they did not.¹⁰⁸ Keegstra was convicted under article 319(2) of Canada’s Criminal Code, which prohibits publicly communicating a statement that ‘willfully promotes hatred against an identifiable group.’¹⁰⁹

Canadian Charter protections, including the right to free speech, can be subject to any ‘reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society.’¹¹⁰ For a restriction on the right to free speech to be permissible under the Canadian Charter, it must meet a two-pronged test: the objective of the restriction must be ‘pressing and substantial’ and it must be proportionate to this objective.¹¹¹ In *Keegstra*, Chief Justice Dickson, writing on behalf of the majority, found article 319(2) to have ‘easily satisfied’ this test, referencing the substantial harm caused by hate, the requirements of Canada’s international treaty obligations, and values recognized in the Charter.

In its judgment, the majority defined hatred as a ‘most extreme emotion that belies reason; an emotion that, if exercised against members of an identifiable group, implies that those individuals are to be despised, scorned, denied respect and made subject to ill-treatment on the basis of group affiliation.’¹¹² It also concluded that the word ‘wilfully’ denoted a ‘stringent standard of *mens rea*’ which ‘is an invaluable means of limiting the incursion of s. 319(2) into the realm of acceptable (though perhaps offensive and controversial) expression.’¹¹³ And the majority disagreed with the submission that section 319(2) was unacceptably broad because it failed to require ‘proof of actual hatred resulting from a communication’, holding that a causal link would be ‘clearly difficult to prove’ and ‘gives insufficient attention to the severe psychological trauma suffered’ by targeted groups.¹¹⁴ The majority also found that the defences allowed under the legislation meant that the danger of overbreadth was ‘significantly reduced’¹¹⁵ and

v. Andrews [1990] 3 S.C.R. 870, 13 December 1990 which also concerned section 319(2) of the Canadian Criminal Code, and *Canada (Human Rights Commission) v. Taylor* [1990] 3 S.C.R. 892, 13 December 1990 on the civil liability provision, s. 13 of the Canadian Human Rights Act.

¹⁰⁸ Canadian Supreme Court, *R v. Keegstra* [1990] 3 S.C.R. 697, 13 December 1990, 714 (Dickson CJ).

¹⁰⁹ Canadian Criminal Code Art. 319(2). See s. II.1.4. (Exclusions, exceptions and defences).

¹¹⁰ Canadian Constitution Act 1982 Part 1 Canadian Charter of Rights and Freedoms.

¹¹¹ Canadian Supreme Court, *R v. Keegstra* [1990] 3 S.C.R. 697, 13 December 1990, 734–735 (Dickson CJ) citing Canadian Supreme Court, *R v. Oakes* [1986] 1 S.C.R. 103.

¹¹² Canadian Supreme Court, *R v. Keegstra* [1990] 3 S.C.R. 697, 13 December 1990, 777 (Dickson CJ). See also T. Mendel, ‘Does International Law Provide for Consistent Rules on Hate Speech?’, in M. Herz & P. Molnar (eds), *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (CUP 2012), 427–428.

¹¹³ Canadian Supreme Court, *R v. Keegstra* [1990] 3 S.C.R. 697, 13 December 1990, 775.

¹¹⁴ *Ibid.*, 776 (Dickson CJ).

¹¹⁵ *Ibid.*, 779 (Dickson CJ): Defences are set out in s. 319(3) and include ‘elements of good faith or honest belief’ and therefore ‘negate directly the mens rea in the offence’, namely if ‘(a) [the defendant] establishes that the statements communicated were true; (b) if, in good faith, he expressed or attempted to establish by argument an opinion on a religious subject; (c) if the statements were relevant to any subject of public interest, the discussion of which was for the public benefit, and if on reasonable grounds he believed them to be true; or (d) if, in good faith, he intended to point out, for the purpose of removal, matters producing or tending to produce feelings of hatred toward an identifiable group in Canada.’

that the impugned provision was a constitutionally valid infringement on the right to free expression.

Accordingly, the majority found article 319(2) was a constitutionally valid infringement on the right to free expression, reasonably justified in a free and democratic society. It had both an important and necessary objective, and was a proportionate means to achieve this objective, in light of Canada's commitment to uphold its foundational values of dignity, equality, and multiculturalism, while combatting the harms of hate, in a narrowly tailored manner offering due process protections.

Three judges of the Supreme Court provided a forceful dissent, arguing that '[t]he breadth of the category of speech [section 319(2)] catches, the absolute nature of the prohibition it applies to such speech, the draconian criminal consequences it imposes coupled with the availability of preferable remedies, and finally, the counterproductive nature of its actual effects ... combine to make it an inappropriate means of protecting our society against the evil of hate propaganda.'¹¹⁶ Justice McLachlin suggested that Canadian courts should 'follow the American approach' and require 'clear and present danger before free speech can be overridden.'¹¹⁷

2.3.2. *United States*

The US Supreme Court's interpretation of the First Amendment, which guarantees that 'Congress shall make no law ... abridging the freedom of speech, or of the press,' has resulted in the United States being one of the most speech-protective jurisdictions across the world.¹¹⁸ First Amendment jurisprudence 'require[s] a concrete link to violence or a breach of the peace' if speech is to be criminalized.¹¹⁹ Criminal penalties for three categories of speech have been found permissible on this basis: speech inciting 'imminent lawless action,' 'fighting words' and 'true threats.'¹²⁰

The most significant category of permissible restrictions to speech is the first one.¹²¹ This was developed in the seminal case of *Brandenburg v. Ohio*, when the Supreme Court ruled that a racist speech by a Ku Klux Klan leader was protected by the First Amendment because the state cannot proscribe advocacy of the use of force 'except where such advocacy is directed to inciting or producing imminent lawless action and is likely to incite or produce such action.'¹²² The Court

¹¹⁶ Ibid., 868. McLachlin J delivered a dissenting judgment, with Sopinka J agreeing and La Forest J agreeing on the issues respecting freedom of expression.

¹¹⁷ Ibid., 807, 822. The dissenting judges criticized the majority's definition of hatred, arguing that 'it is not only the breadth of the term 'hatred' which presents dangers; it is its subjectivity.'

¹¹⁸ Cf., ch. 5 (Espionage and Official Secrets Laws), s. I.3.4.1. (United States: Pentagon Papers); ch. 6 (Terrorism Laws), s. I.2.4.1. (United States). A number of sources chart the US Supreme Court's historical path towards this more speech protective position: see, e.g., M. Rosenfeld, 'Hate Speech in Constitutional Jurisprudence: A Comparative Analysis', in M. Herz & P. Molnar (eds), *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (CUP 2012); E. Bleich, 'Freedom of Expression versus Racist Hate Speech: Explaining Differences Between High Court Regulation in the USA and Europe' (2014) 40 *Journal of Ethnic and Migration Studies* 283.

¹¹⁹ A. Clooney & P. Webb, 'The Right to Insult in International Law' (2017) 48(2) *Columbia Human Rights Law Review* 1, 50.

¹²⁰ Ibid., 48.

¹²¹ US Supreme Court, *Brandenburg v. Ohio* 395 U.S. 444, 9 June 1969.

¹²² Ibid.

concluded that in that case ‘the Klan may have *advocated* violence but had not *incited* it’.¹²³

Subsequent US hate speech jurisprudence has solidified the requirement that speech must have a clear call to violence or lawless action if it is to be criminalized.¹²⁴ For example, when an anti-war protester said ‘we’ll take the fucking street later’ in the context of a 100–150 person demonstration that had moved onto the public street and blocked traffic, the Supreme Court held that ‘at worst, it amounted to nothing more than advocacy of illegal action at some indefinite future time’ and that the words ‘could not be punished by the State on the ground that they had a “tendency to lead to violence”’.¹²⁵ And in considering whether a speech made at a civil rights meeting in Mississippi that advocated for a boycott of white merchants was protected by *Brandenburg v. Ohio*, the Supreme Court held that ‘an advocate must be free to stimulate his audience with spontaneous and emotional appeals for unity and action,’ and there was no evidence that the speaker ‘authorized, ratified or directly threatened acts of violence.’¹²⁶ So in both cases the Court found that the First Amendment protected the impugned speech.

The second category of permissible restrictions to speech pertains to ‘true threats,’ meaning statements through which ‘the speaker means to communicate a serious expression of an intent to commit an act of unlawful violence to a particular individual or group of individuals.’¹²⁷ The statement at a public rally that ‘[i]f they ever make me carry a rifle the first man I want to get in my sights is L.B.J.’ was found to be ‘political hyperbole,’ rather than such a true threat.¹²⁸ Although scholars have questioned the parameters of this category of unprotected speech,¹²⁹ the Supreme Court recently confirmed that a defendant can be convicted under federal statutes criminalizing threats of violence.¹³⁰ In his partial concurrence, Justice Alito stated that ‘it is settled that the Constitution does not protect true threats,’ because such statements ‘inflict great harm and have little if any social value. A threat may cause serious emotional stress for the person threatened and those who care about that person, and a threat may lead to violent confrontation.’¹³¹

The final category of speech not protected by the First Amendment is that of ‘fighting words,’ which encompasses ‘those which, by their very utterance, inflict injury or tend to incite an immediate breach of the peace.’¹³² This doctrine was enunciated in *Chaplinsky v. New Hampshire*, some 27 years before the Court decided *Brandenburg*. A conviction

¹²³ Rosenfeld (n 118) 254.

¹²⁴ See, e.g., US Supreme Court, *NAACP v. Claiborne Hardware Co.* 458 U.S. 886, 2 July 1982, 928.

¹²⁵ US Supreme Court, *Hess v. Indiana* 414 U.S. 105, 19 November 1973, 108–109.

¹²⁶ US Supreme Court, *NAACP v. Claiborne Hardware Co.* 458 U.S. 886, 2 July 1982, 928–930. Cf. ch. 6 (Terrorism Laws), s. I.2.4.1. (United States).

¹²⁷ US Supreme Court, *Virginia v. Black* 538 U.S. 343, 7 April 2003, 359. See also US Supreme Court, *Counterman v. Colorado* 600 U.S. 66, 27 June 2023.

¹²⁸ US Supreme Court, *Watts v. United States* 394 U.S. 705, 21 April 1969, 708.

¹²⁹ See, e.g., P. Crane, ‘True Threats’ and the Issue of Intent’ (2006) 92 *Virginia Law Review* 1225; S. Gay, ‘A Few Questions About Cross Burning, Intimidation and Free Speech’ (2005) 80 *Notre Dame Law Review* 1287; ‘Elonis v. United States’ (2015) 129 *Harvard Law Review* 331.

¹³⁰ US Supreme Court, *Elonis v. United States* 575 U.S. 723, 1 June 2015.

¹³¹ *Ibid.*, Alito J, concurring in part and dissenting in part.

¹³² US Supreme Court, *Chaplinsky v. New Hampshire* 315 U.S. 568, 9 March 1942, 571–572.

for saying the words ‘You are a God damned racketeer’ and a ‘damned Fascist’ to a public official was upheld on the basis that such words were ‘likely to provoke the average person to retaliation, and thereby cause a breach of the peace.’¹³³ However, this doctrine has been described by the Third Circuit as an ‘extremely narrow one’, and by commentators as having emerged to be ‘strikingly similar to the clear and present danger test’ (later replaced by the ‘imminent lawless action’ test in *Brandenburg*).¹³⁴ For example, the Supreme Court considered that wearing a jacket that said ‘Fuck the Draft’ did not amount to ‘fighting words’, stating that the doctrine only covered ‘personally abusive epithets which, when addressed to the ordinary citizen, are, as a matter of common knowledge, inherently likely to provoke violent reaction.’¹³⁵ And in *Gooding v. Wilson*, where the impugned speech included the phrases ‘You son of a bitch, I’ll choke you to death’, the Court found that a statute criminalizing the use of ‘opprobrious words or abusive language, tending to cause a breach of the peace’ to be overbroad because it was not limited to words that ‘have a direct tendency to cause acts of violence by the person to whom, individually, the remark was addressed.’¹³⁶

Beyond these three categories and their requisite requirements for a clear link between speech and violent or illegal acts, US First Amendment jurisprudence also clearly rejects content-based restrictions to speech. For example, when a defendant was convicted for burning a cross pursuant to an ordinance that prohibited displaying symbols which knowingly ‘arouse[] anger, alarm or resentment in others on the basis of race, color, creed, religion or gender’, the Supreme Court held that the relevant statute was unconstitutional ‘in that it prohibits otherwise permitted speech solely on the basis of the subjects the speech addresses.’¹³⁷ And in a more recent case, the Court confirmed that speech that demeans ‘on the basis of race, ethnicity, gender, religion, age, disability, or any other similar ground is hateful; but the proudest boast of our free speech jurisprudence is that we protect the freedom to express “the thought that we hate”’.¹³⁸

A recent example of a case in which the high bar provided by the First Amendment was met was the criminal conviction and sentence of Ryder Winegar. He had left six

¹³³ *Ibid.*, 574.

¹³⁴ See US Court of Appeals, *Johnson v. Campbell* 332 F.3d 199, 5 June 2003, 212; M. Mannheim, ‘The Fighting Words Doctrine’ (1993) 3 *Columbia Law Review* 1527; Clooney & Webb, ‘The Right to Insult in International Law’ (n 119) 48 (noting that the category ‘has not been heavily relied on in practice and its contours are very narrow’); ‘The Demise of the “Chaplinsky” Fighting Words Doctrine: An Argument for Its Interment’ (1993) 106 *Harvard Law Review* 1129.

¹³⁵ US Supreme Court, *Cohen v. California* 403 U.S. 15, 7 June 1971, 20.

¹³⁶ US Supreme Court, *Gooding v. Wilson* 405 U.S. 518, 23 March 1972, 523 (‘Our decisions since *Chaplinsky* have continued to recognize state power constitutionally to punish “fighting” words under carefully drawn statutes not also susceptible of application to protected expression’).

¹³⁷ US Supreme Court, *R.A. V. v. City of St Paul, Minnesota* 505 U.S. 377, 22 June 1992, 381. See also US Supreme Court, *Matal v. Tam*, 137 S. Ct. 1744, 19 June 2017, 1751, Opinion of Alito J, joined by Roberts CJ, and Thomas and Breyer JJ (finding that an act prohibiting the registration of trademarks that may ‘disparage ... or bring ... into contempt or disrepute’ any living persons to be unconstitutional, on the basis that speech ‘may not be banned on the ground that it expresses ideas that offend’). (‘From 1791 to the present ... our society, like other free but civilized societies, has permitted restrictions upon the content of speech in a few limited areas, which are ‘of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality’: US Supreme Court, *R.A. V. v. City of St Paul, Minnesota* 505 U.S. 377, 22 June 1992, 382–383).

¹³⁸ US Supreme Court, *Matal v. Tam* 137 S. Ct. 1744, 19 June 2017, 1764, Opinion of Alito J, joined by Roberts CJ, and Thomas and Breyer JJ.

members of Congress threatening voicemails two days after President Biden's election was confirmed by the electoral college in December 2020. One message described the 'massive fraud' that had taken place in the United States, and threatened 'if you don't support it, we're going to drag you out and we're going to hang you by your neck to die'. Winegar was sentenced to 33 months in federal prison after pleading guilty to six counts of threatening members of Congress and one count of transmitting interstate threatening communications.¹³⁹

2.4. Middle East and Africa

2.4.1. Tunisia

An example of the misuse of hate speech laws in a religious context is the conviction of Emna Chargui, a 27-year-old Tunisian blogger, for 'inciting hatred between religions through hostile means or violence' and 'infringing an authorized religion' pursuant to Tunisia's 'Law on Freedom of the Press'.¹⁴⁰ In 2020, Chargui shared text that imitated the format of a Quranic paragraph ('sura'), titled 'Sura Corona', encouraging readers to follow science and wash their hands, stating 'there's no difference between kings and slaves, follow science and ignore traditions'.¹⁴¹ The verse made no comment on religion, beyond using the style of a sura. Chargui did not even write the text but reposted it because she 'thought it was a good way to make people be . . . careful with the coronavirus, with a style that everyone knows'.¹⁴² At trial, prosecutors failed to present any evidence showing that Chargui had an intention to incite hostility or violence against Muslims, or that such violence was likely to occur as a result of her post. Instead, the prosecution argued that posting 'text similar to the Qu'ran' constituted an offence under articles 52 and 53 of Tunisia's press law. Article 52 sanctions the act of calling directly to hatred between races, religions or the population by inciting discrimination and using hostile means, violence or dissemination of ideas based on racial discrimination. Article 53 provides for the sentence of imprisonment and a fine for the act of intentionally infringing on a recognized religious rite.¹⁴³ She was convicted and sentenced to six months' imprisonment and a fine.¹⁴⁴ The decision has been condemned by human rights groups, which have described it as a 'bitter blow to freedom of expression in Tunisia'.¹⁴⁵ Although reports had suggested that Chargui's lawyers intended to appeal the decision, there is no record of the appeal and Chargui has since left the country.¹⁴⁶

¹³⁹ G. Harkins, 'A man said he'd hang 6 members of Congress who didn't "get behind" Trump, feds say. He got 33 months in prison' (The Washington Post, 2 December 2021).

¹⁴⁰ See CFJ, 'Statement on the Conviction of Emna Chargui in Tunisia' (15 July 2020); A. Benchemsi & N. Slama, 'Humor Comes at a Price in Morocco and Tunisia' (Human Rights Watch, 13 May 2020).

¹⁴¹ 'Tunisia: Woman gets six-month jail sentence for coronavirus post mimicking Quran' (Middle East Eye, 15 July 2020).

¹⁴² L. Blaise & E. Peltier, 'Tunisian Woman Sentenced to Prison Over Joke Alluding to the Quran' (The New York Times, 17 July 2020).

¹⁴³ Tunisian Decree No. 115 on Freedom of the Press, Printing, and Publishing of 2 November 2011 Arts 52–53.

¹⁴⁴ Blaise & Peltier (n 142).

¹⁴⁵ Amnesty International, 'Tunisia: Blogger Emna Chargui sentenced to six months in prison for social media post' (15 July 2020).

¹⁴⁶ Human Rights Watch, 'Tunisia, Events of 2020'.

2.4.2. Saudi Arabia

In Saudi Arabia, prosecutions for anti-religious speech are routine and often accompanied by harsh penalties. Such was the case of Raif Badawi, who was sentenced to 10 years' imprisonment and one thousand lashes for insulting Islam.¹⁴⁷ In 2006, Badawi launched an 'online forum for political and social debate' called 'Saudi Arabian Liberals' and was subsequently arrested. He was detained for almost 10 years and remains subject to a travel ban.¹⁴⁸

The criminal proceedings against Badawi have been described by the UN as a 'flagrant misadministration' of justice 'manifested in the repetition of ... trials and convictions at different levels of the courts.'¹⁴⁹ He was ultimately sentenced to five years' imprisonment for contempt of religion, insulting Islam and disseminating 'electronic materials that may undermine public order or mock public morality or national figures' on the basis of pro-secular speech on his blog. This included posts which, according to the prosecutor, questioned the 'validity of prayer; wondered why Valentine's day is not celebrated in Saudi Arabia; and mocked ... religious and political figures.'¹⁵⁰ However, it has been said that through his blog posts Badawi 'unmasked a culture of corruption and criminality, as well as the impunity that underpinned them. He challenged religious intolerance and extremism, and sparked a discussion on modernization' in Saudi Arabia.¹⁵¹ The judge added two years to his sentence for 'insulting Islam and Saudi's religious police' during televised interviews and three months for 'parental disobedience', resulting in a total sentence of seven years and 600 lashes.¹⁵² On appeal, the sentence was increased to 10 years' imprisonment, a fine amounting to approximately \$250,000 US dollars and a 10-year ban on travel and journalistic activity. The appellate court rejected Badawi's submission that he should not be held responsible for speech by third parties posted on the blog on the basis that he was in control of it and 'disregarded the blasphemy being spread through it.'¹⁵³

III. International Legal Standards

Freedom of expression is enshrined in article 19 of the Universal Declaration of Human Rights, article 19 of the ICCPR and regional human rights instruments including the European Convention, American Convention, African Charter on Human and Peoples' Rights, Arab Charter on Human Rights and the ASEAN Human Rights Declaration.¹⁵⁴

¹⁴⁷ See, e.g., WGAD, *Sheikh Suliaman al-Rashudi and others v. Saudi Arabia* (Opinion no. 38/2015), 4 September 2015, §§24–27 and Raoul Wallenberg Centre for Human Rights, 'Appeal for Clemency for Raif Badawi'.

¹⁴⁸ WGAD, *Sheikh Suliaman al-Rashudi and others v. Saudi Arabia* (Opinion no. 38/2015), 26 October 2015, §24; 'Raif Badawi: Saudi blogger freed after decade in prison' (BBC, 11 March 2022).

¹⁴⁹ WGAD, *Sheikh Suliaman al-Rashudi and others v. Saudi Arabia* (Opinion no. 38/2015), 26 October 2015, §80.

¹⁵⁰ Columbia Global Freedom of Expression Database, 'Saudi Public Prosecutor v. Raif Badawi: He was initially charged with the capital crime of apostasy but this was later dropped.'

¹⁵¹ B. Silver & E. Abitbol, 'One Way Crown Prince Mohamed bin Salman Can Prove He Is Sincere About His Reforms: Free Raif Badawi' (TIME, 5 April 2018).

¹⁵² Human Rights Watch, 'Saudi Arabia: 600 Lashes, 7 Years for Activist' (30 July 2013).

¹⁵³ Jeddah District Appeal Court decision dated 27 May 2014 (unofficial translation).

¹⁵⁴ See ECHR Art. 10; ACHR Art. 13; ACHPR Art. 9; Arab Charter Art. 32; ASEAN Human Rights Declaration Art. 23; and Cairo Declaration on Human Rights in Islam Art. 22. Article 32 of the Arab Charter provides the same

Each defines the right to freedom of expression broadly. But international law also provides that states may—and in certain circumstances *must*—punish hate speech. This *requirement* that certain speech should be penalized, and even in some instances criminalized, is a key difference between international standards governing hate speech laws and standards relating to other types of speech. For instance, although speech that is defamatory in nature may be penalized, there is no international law *obligation* to do so.¹⁵⁵

In the case of hate speech, three UN treaties mandate the imposition of penalties. First, article 20 of the ICCPR requires that ‘propaganda for war’ and ‘advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence . . . *shall be prohibited by law*’. Second, article 4 of CERD requires penalties for racist speech. And, third, the Genocide Convention requires penalties for speech that incites genocide. The ICCPR provision allows for civil and administrative penalties, as well as potentially criminal ones,¹⁵⁶ while the other two provisions mandate penalties that are criminal in nature.

In addition to mandatory restrictions on speech, international human rights law provides guidance on when such restrictions are permissible. Under article 19(3) of the ICCPR, a state is permitted to penalize speech if the state can prove that the restriction is (1) ‘provided by law’ and (2) necessary to (3) protect one of the following legitimate objectives: ‘the rights or reputations of others’, ‘national security’, ‘public order’ or ‘public health or morals.’ Mandatory restrictions on speech must also comply with this three-part test.¹⁵⁷

1. Mandatory Restrictions on Hate Speech

Article 20 of the ICCPR mandates that ‘propaganda for war’ and ‘advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or

legitimate aims by which expression can be restricted as article 19(3) of the ICCPR, but adds a caveat that rights must be ‘exercised in conformity with the fundamental values of society’, and the Charter currently lacks an enforcement mechanism. The ASEAN Human Rights Declaration is a non-binding instrument which, although providing a broad right to freedom of opinion and expression in article 23, has a number of general limitation clauses, for example article 7, which stated that ‘realisation of human rights must be considered in the regional and national context bearing in mind different political, economic, legal, social, cultural and religious backgrounds’. The Cairo Declaration, a non-binding instrument, provides more extensive limits to speech than article 19 of the ICCPR, but the Organization of Islamic Cooperation has been revising the instrument, with finalization of an amended Declaration on Human Rights postponed due to the COVID-19 pandemic: see T. Kayaoglu, ‘The Organization of Islamic Cooperation’s declaration on human rights: Promises and pitfalls’ (Brookings, 28 September 2020).

¹⁵⁵ Although an important component of the practical reality of litigating hate speech cases, due process issues such as the burden and standard of proof and limitation periods are not within the scope of this chapter. See A. Clooney & P. Webb, *The Right to a Fair Trial in International Law* (OUP 2020).

¹⁵⁶ The Human Rights Committee has not, to date, stated that any particular speech requires criminalization under Art. 20 of the ICCPR. Cf. HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996 (overruled) (where the Committee found the French Gayssot Act which required criminalization of speech to be compliant with the ICCPR).

¹⁵⁷ HRC, General Comment No. 34 (2011), §50, citing HRC, *Ross v. Canada* (Comm. no. 736/1997), 18 October 2000. See also HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, §10.4.

violence ... *shall be* prohibited by law'. Speech falling within article 20 *must* be legally prohibited, but criminal sanctions are not required.

CERD and the Genocide Convention, on the other hand, mandate *criminal* penalties for certain forms of hate speech. CERD requires in article 4 that states '[s]hall declare [as] an offence ... all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all ... incitement to ... acts [of violence] against any race or group of persons of another colour or ethnic origin'. This must, however, be done 'with due regard' to other human rights principles and obligations, including freedom of expression.¹⁵⁸ In its General Recommendation 35, the UN Committee on the Elimination of all Forms of Racial Discrimination, which is established by the Convention to monitor implementation of the treaty, clarified that 'states parties should take into account ... the intention of the speaker' and 'the imminent risk or likelihood that the conduct desired or intended by the speaker will result from the speech in question' when sanctioning racist hate speech that falls within the scope of article 4.¹⁵⁹ This authoritative guidance, as well as the 'due regard' clause, therefore significantly narrow the scope of the mandate to criminalize speech.

Article 3 of the Genocide Convention requires that 'direct and public incitement to commit genocide' *shall be* punishable. This has been interpreted by the International Criminal Tribunal for Rwanda as meaning punishment through criminal law.¹⁶⁰ The elements of this crime include the need to prove that: (1) the speaker specifically intended to cause genocide through their speech; (2) the incitement was 'direct', based on an assessment of the local context to determine how the speech would be understood by the intended audience; (3) the intended harm was genocide, which means a specific intent to eradicate particular groups in whole or in part through certain acts; and (4) the statements were made in public.¹⁶¹

The American Convention on Human Rights is the only regional human rights treaty that mandates the criminal prohibition of hate speech. This mandatory prohibition, contained in article 13 of the Convention, provides that '[a]ny propaganda for war and any advocacy of national, racial, or religious hatred that constitute incitements to lawless violence or to any other similar 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'. It is therefore narrower than article 20 of the ICCPR and is limited to situations of (1) 'propaganda', (2) that incites (3) an objective and imminent risk of 'lawless violence' or 'similar' harm (4) against any person

¹⁵⁸ Article 4 provides that states must have 'due regard to the principles embodied in the Universal Declaration of Human Rights and the rights expressly set forth in article 5 of this Convention'. Article 19 of the UDHR and subparagraph (d)(viii) of CERD article 5 protect freedom of expression.

¹⁵⁹ CERD Committee, General Recommendation No. 35 (2013), §16.

¹⁶⁰ See ICTR, *Prosecutor v. Akayesu* (ICTR-96-4-T), Judgement, 2 September 1998, §551 ('At the time the Convention on Genocide was adopted, the delegates agreed to expressly spell out direct and public incitement to commit genocide as a specific crime'). See also ICTR, *Prosecutor v. Nahimana et al.* (ICTR-99-52-A), Judgement, 28 November 2007, §678; UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §40.

¹⁶¹ ICTR, *Prosecutor v. Nahimana et al.* (ICTR-99-52-A), Judgement, 28 November 2007, §§698–701, 709.

or group of persons on any grounds including those of race, colour, religion, language or national origin.

1.1. ICCPR Article 20

Article 20 of the ICCPR was first proposed by a Soviet diplomat who considered that such a provision would be a ‘powerful weapon . . . to restrict the dissemination of Nazi-Fascist propaganda.’¹⁶² However, a number of delegations drafting the treaty opposed the proposal, with the US delegation warning that ‘any criticism of public or religious authorities might all too easily be described as incitement to hatred and consequently prohibited’; and the UK delegates arguing that hatred was ‘not easy to define as a penal offence.’¹⁶³

The UN Human Rights Committee has made clear that restrictions imposed under article 20 of the ICCPR must also meet article 19’s tripartite test of legality, legitimacy and necessity.¹⁶⁴ The Human Rights Committee has not provided a detailed interpretation of article 20¹⁶⁵ but in the *Rabbae v. the Netherlands* case it closely reviewed the application of hate speech provisions enacted to give effect to article 20 in Dutch law and in doing so provided important guidance.¹⁶⁶

The *Rabbae* case was brought by Dutch-Moroccan nationals who claimed that the Netherlands had violated article 20 by failing to convict Geert Wilders, a right-wing Dutch politician, for speech that was discriminatory on the grounds of race and religion.¹⁶⁷ The Committee found that the Netherlands had not violated the ICCPR since the treaty ‘does not provide individuals with a right to have [specific] individuals prosecuted’ let alone ‘ensure that a person who is charged with incitement to discrimination, hostility and violence will invariably be convicted.’¹⁶⁸

Although the Committee has not provided definitions of the terms ‘advocacy’, ‘hatred’ or ‘incitement’, members of the Committee were satisfied that the definition of incitement under Dutch law—‘inflammatory behaviour that incites the commission of criminal offences or acts of violence’—was a proper implementation of article 20(2).¹⁶⁹ The Committee also confirmed that the prohibition provided by article 20 does not mandate criminalization of speech, but can be fulfilled by civil and administrative penalties.¹⁷⁰ A concurring opinion in *Rabbae* also noted that article 20 sets ‘a high bar’ and

¹⁶² See Clooney & Webb, ‘The Right to Insult in International Law’ (n 119) 17.

¹⁶³ *Ibid.*; see ch. 1 (Introduction), s. II.1.2.1.2. (ICCPR Article 20).

¹⁶⁴ HRC, General Comment No. 34 (2011), §§50–52. See s. III.2. (Discretionary Restrictions on Hate Speech).

¹⁶⁵ In 1983, the Human Rights Committee issued a General Comment that focused primarily on improved reporting in states parties’ periodic reports. HRC, General Comment No. 10 (1983), §1.

¹⁶⁶ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016; see s. II.2.1.1. (The Netherlands). Prior to *Rabbae*, the Committee had referenced and in one instance found a violation of article 20, but without opining on its parameters. See HRC, *J. R. T. and the W. G. Party v. Canada* (Comm. no. 104/1981), 6 April 1983, §8(b); HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996, Individual Opinion of Committee members Elizabeth Evatt and David Kretzmer, co-signed by Eckart Klein, §4; HRC, *Ross v. Canada* (Comm. no. 736/1997), 18 October 2000, §11.5.

¹⁶⁷ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, §3.1.

¹⁶⁸ *Ibid.*, §§10.3, 10.7, 11.

¹⁶⁹ *Ibid.*, §10.7, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §6.

¹⁷⁰ *Ibid.*, §10.4. This conclusion has also been reached by the UN Special Rapporteur: UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §47.

does not require legal prohibition of all advocacy of national, racial or religious hatred, ‘but only of such advocacy that also “constitutes *incitement* to discrimination, hostility or violence”’.¹⁷¹

Further guidance about the parameters of article 20 is provided by ‘soft law’ instruments such as the Rabat Plan of Action, a codification of international standards endorsed by the UN High Commissioner for Human Rights.¹⁷² These standards conclude that ‘a high threshold’ should apply to ‘the application of article 20’.¹⁷³ They also recognize that ‘[c]riminal sanctions related to unlawful forms of expression’ should only apply in the most serious cases.¹⁷⁴

The Rabat Plan of Action identifies six factors relevant to determining whether hateful speech can reach this high threshold and be criminalized. These are: (1) the context in which the statement was made; (2) the position or status of the speaker; (3) the intent of the speaker (noting that negligence and recklessness would not suffice); (4) the content and form of the speech; (5) the ‘reach’ of the speech including the nature and size of the audience; and (6) the likelihood of imminent harm arising from the speech.¹⁷⁵ These principles have been endorsed by the UN High Commissioner for Human Rights and the UN Special Rapporteur on Freedom of Opinion and Expression and they remain relevant to assessing the validity of hate speech laws that seek to give effect to article 20.¹⁷⁶

In addition, the UN Special Rapporteur stated that the plain language of article 20 establishes three threshold elements that a state must demonstrate for speech to qualify for prohibition under this provision:¹⁷⁷ (1) ‘advocacy’ of national, racial or religious hatred should mean the ‘explicit, intentional, public and active support and promotion of hatred towards the target group’;¹⁷⁸ (2) hatred must refer to the ‘intense and irrational emotions of opprobrium, enmity and detestation’; and (3) there must be ‘incitement’ to the harms listed in article 20, meaning that the speech should ‘create an *imminent risk*’ of such harm.¹⁷⁹ He also noted that ‘hostility’ should be interpreted to mean ‘a manifestation of hatred beyond a mere state of mind’.¹⁸⁰

¹⁷¹ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §4.

¹⁷² Rabat Plan of Action on the prohibition of advocacy of nation, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.

¹⁷³ *Ibid.*, §29.

¹⁷⁴ *Ibid.*, §34. This conclusion has also been reached by the UN Special Rapporteur: UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §47.

¹⁷⁵ Rabat Plan of Action, §29.

¹⁷⁶ UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §57(a). See s. V. (Approach of Private Companies to Online Hate Speech).

¹⁷⁷ UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §43.

¹⁷⁸ *Ibid.*, §44(b) citing principle 12.1 of the Camden Principles (advocacy ‘is to be understood as requiring an intention to promote hatred publicly towards the target group’); Rabat Plan of Action, n 5, §29(c).

¹⁷⁹ UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §44(a), (c) citing principle 12.1 of the Camden Principles (emphasis added). See also Rabat Plan of Action, §29.

¹⁸⁰ UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §44(e).

1.2. CERD Article 4

The CERD treaty also requires criminal penalties for certain hate speech. More specifically article 4 provides that states ‘[s]hall declare [as] an offence ... all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all ... incitement to ... acts [of violence] against any race or group of persons of another colour or ethnic origin.’ This must, however, be done ‘with due regard’ to other human rights principles and obligations, including freedom of expression.¹⁸¹ Article 4 refers to ‘offense[s] punishable by law’, and thus explicitly mandates criminal sanctions.

As a result of the breadth of the wording of article 4(a), a number of states entered reservations and declarations that seek to limit its scope and provide more protection for speech. The United States, for example, entered a reservation that it ‘does not accept any obligation under this Convention’, in particular under article 4, to restrict protections of freedom of speech, expression and association ‘to the extent that they are protected by the Constitution and laws of the United States.’¹⁸² Similar reservations were made by some European states as well as Japan, Guyana and Jamaica.¹⁸³

However, in 2013 the CERD Committee—the expert body charged with monitoring implementation of the Racial Discrimination Convention—issued ‘General Recommendation 35’, which adopted a narrower interpretation of the scope of article 4 than the Committee had previously espoused.¹⁸⁴ First, the Committee noted the five contextual factors ‘adapted from’ the Rabat Plan of Action that ‘should be taken into account’ when determining if speech should be punishable under article 4. These are: the ‘content and form of speech’; the ‘economic, social and political climate’ at the time speech was disseminated; the ‘position or status of the speaker’ and the audience to which the speech was directed; ‘the reach of the speech’ and its ‘objectives’.¹⁸⁵ Secondly, the Committee clarified that in interpreting and applying the term ‘incitement’, states parties should consider ‘the intention of the speaker, and the imminent risk or likelihood that the conduct desired or intended by the speaker will result from the speech in question.’¹⁸⁶

Indeed the Committee went further and clarified that the issue of intent and imminence should be relevant not only to incitement but to *all* speech criminalized by article 4, including ‘all dissemination of ideas based on racial superiority or hatred’, thereby interpreting this provision more closely in line with article 20 of the ICCPR.¹⁸⁷ As a result, the lead drafter of General Recommendation 35 considered that the Recommendation ‘decisively rejects any suggestion of a “strict liability” approach to

¹⁸¹ CERD Art. 4(a). Article 4 provides that states must have ‘due regard to the principles embodied in the Universal Declaration of Human Rights and the rights expressly set forth in article 5 of this Convention’. Article 19 of the UDHR and sub-paragraph (d)(viii) of CERD article 5 protect freedom of expression.

¹⁸² UNTC Depository, Status of Treaties, ch. IV.2 CERD, Declarations and Reservations, United States of America.

¹⁸³ See ch. 1 (Introduction), s. II.1.3.1.3. (CERD Article 4) setting out reservations by Guyana, Jamaica and Japan as well as Austria, Belgium, France, Ireland, Italy and Monaco.

¹⁸⁴ CERD Committee, General Recommendation No. 35 (2013).

¹⁸⁵ *Ibid.*, §15, n 17.

¹⁸⁶ *Ibid.*, §16.

¹⁸⁷ *Ibid.*, §§13, 16.

dissemination and incitement ... [by linking] them with principles of criminal law on mental elements in crime.¹⁸⁸

The Committee also clarified that the ‘due regard’ language of article 4 requires ‘strict compliance with freedom of expression guarantees’,¹⁸⁹ and that any criminal restrictions on speech ‘should be reserved for serious cases, to be proven beyond reasonable doubt’, and ‘governed by principles of legality, proportionality and necessity’.¹⁹⁰ Finally, the Committee emphasized that ‘the expression of ideas and opinions made in the context of academic debates, political engagement or similar activity, and without incitement to hatred contempt, violence or discrimination, should be regarded as legitimate exercises of the right to freedom of expression’.¹⁹¹

This clarification of the scope of article 4 brings the CERD Committee’s pronouncements closer to the approach adopted by the Human Rights Committee.¹⁹² And prior practice that did not consistently require specific intent and incitement, or address the applicability of ICCPR article 19(3)’s legality, necessity and proportionality principles in hate speech cases, is likely to be considered overturned.¹⁹³

The Committee did, however, also take a step in the opposite direction by approving an expansion of potential liability in the definition of harms. First, it found that incitement to ‘contempt’ can be a harm justifying criminalization under article 4,¹⁹⁴ despite the fact that the text refers only to ‘hatred’ ‘discrimination’ and ‘violence’.¹⁹⁵ The Committee also found that although the title of the Convention suggests that it is limited to ‘racial’ hatred, the harms that are recognized also extend to ‘colour, descent or national or ethnic origin’.¹⁹⁶

1.3. Genocide Convention Article 3

The most extreme form of hate speech addressed by international human rights law is ‘direct and public incitement to commit genocide’. States are required, under the Genocide Convention, to punish such speech through the criminal law.¹⁹⁷

¹⁸⁸ P. Thornberry, ‘International Convention on the Elimination of All Forms of Racial Discrimination: The Prohibition of “Racist Hate Speech”’, in T. McGonagle & Y. Donders (eds), *The United Nations and Freedom of Expression and Information: Critical Perspectives* (CUP 2015), 121, 131. Thornberry was a member of the Committee from 2001–2014.

¹⁸⁹ CERD Committee, General Recommendation No. 35 (2013), §19. See also UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §15.

¹⁹⁰ CERD Committee, General Recommendation No. 35 (2013), §12 (citing to HRC, General Comment No. 34 (2011)).

¹⁹¹ CERD Committee, General Recommendation No. 35 (2013), §25.

¹⁹² See P. Thornberry, *The International Convention on the Elimination of All Forms of Racial Discrimination* (OUP 2016), 297–298, 301–302 (‘[T]he fresh reading of Article 4 takes the Convention closer to the ICCPR’); UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §15 (noting the CERD Committee’s ‘converging interpretations’ with the HRC).

¹⁹³ See, e.g., CERD, *Adan v. Denmark* (Comm. no. 43/2008), 13 August 2010, §§7.2, 7.7; CERD, *The Jewish Community of Oslo v. Norway* (Comm. no. 30/2003), 15 August 2005, §§2.5, 10.5.

¹⁹⁴ CERD Committee, General Recommendation No. 35 (2013), §§13, 25.

¹⁹⁵ CERD Art. 4(a), CERD Committee, General Recommendation No. 35 (2013), §13(b), (d).

¹⁹⁶ CERD Committee, General Recommendation No. 35 (2013), §§6, 13. See also CERD Art. 1 which explicitly states: ‘In this Convention, the term “racial discrimination” shall mean any distinction, exclusion, restriction or preference based on race, colour, descent, or national or ethnic origin.’

¹⁹⁷ Genocide Convention Art. 3(c). The Rome Statute of the ICC criminalizes incitement to genocide in a different format: ‘the crime of genocide’ is listed as one of the crimes within the jurisdiction of the Court under Art. 5 and is defined under Art. 6 without reference to direct and public incitement to commit genocide. However, Art.

Genocide offences have a high bar: the International Court of Justice has held that any crime under article 3 of the Convention, including incitement to commit genocide, must be proven by evidence that is ‘fully conclusive.’¹⁹⁸ The key elements of the crime are set out in the *Nahimana* decision of the ICTR¹⁹⁹ in which three Rwandan media workers were convicted of direct and public incitement to genocide for disseminating extremist Hutu propaganda through radio broadcasts and newspapers during the Rwandan genocide.²⁰⁰

In this case—known as the Media Trial—the ICTR Appeals Chamber held that in order for speech to incite genocide, the speech must be public,²⁰¹ the speaker must have ‘the intent to directly and publicly incite others’ to commit this crime,²⁰² and the speech must be ‘direct,’ based on ‘the meaning of the words used in the specific context.’ This means that if the relevant words do not amount to an explicit appeal to commit genocide, it must be shown that the intended audience ‘immediately grasped’ their implication.²⁰³

The *Nahimana* decision distinguished ‘mere’ hate speech from incitement to commit genocide, recognizing that more stringent requirements applied to the latter.²⁰⁴ Judge Meron also emphasized in his dissent that ‘there is no settled norm of customary

25 provides for modes of individual criminal responsibility across the Statute and specifically notes ‘in respect of the crime of genocide, directly and publicly incites others to commit genocide’: Art. 25(3)(e). Commentators have described this as having ‘downgraded’ incitement ‘from a standalone crime in itself to a mode of participation in genocide’ and that as ‘proof of causation is required for modes of liability for completed crimes and furnishing such proof is an arduous undertaking, the ICC’s unusual formulation of incitement to genocide means that prosecutions at the ICC for the crime are less likely to succeed’: R.A. Wilson, *Incitement on Trial: Prosecuting International Speech Crimes* (CUP 2017), 34.

¹⁹⁸ ICJ, *Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Croatia v. Serbia)*, Judgment of 3 February 2015, §178; see also §407 (when direct evidence of specific intent is absent, the Court requires that ‘the only reasonable conclusion to be drawn [from conduct] is an intent’ to ‘destroy that substantial part of the group’).

¹⁹⁹ ICTR, *Prosecutor v. Nahimana et al.* (ICTR-99-52-A), Judgement, 28 November 2007, §677.

²⁰⁰ The defendants were convicted under article 2(3)(c) of the Statute of the International Tribunal for Rwanda, which prohibits incitement to genocide in the same terms as the Genocide Convention. Two of them were political figures and the founders of an independent radio station Radio Télévision Libre des Mille Collines (RTLM) and the third was the owner and editor of a newspaper.

²⁰¹ In subsequent cases the ICTR elaborated on the requirement that the speech be ‘public’ to constitute ‘incitement’. In one case the tribunal noted that ‘all convictions before the Tribunal for direct and public incitement to commit genocide involve speeches made to large, fully public assemblies, messages disseminated by the media, and communications made through a public address system over a broad public area’. Crowds of between 20–30 people have satisfied this requirement and convictions may be based on ‘communications to smaller audiences when the incriminating message is given in a public space to an unselected audience’. However, a meeting limited to only a group of public officials, even though it was in the ‘presence of a journalist’, did not fulfil the ‘public’ element of incitement. See ICTR, *Prosecutor v. Nzabonimana* (ICTR-98-44D-T), Judgement and Sentence, 31 May 2012, §§1754, 1772; ICTR, *Nzabonimana v. Prosecutor* (ICTR-98-44D-A), Judgement, 29 September 2014, §§126, 231, 385. See s. II.2.1.1. (The Netherlands).

²⁰² ICTR, *Prosecutor v. Nahimana et al.* (ICTR-99-52-A), Judgement, 28 November 2007, §677; see also §709 (When determining intent, ‘the fact that a speech leads to acts of genocide could be an indication that in that particular context the speech was understood to be an incitement to commit genocide’, but this is not the only way to prove intent). The Tribunal also confirmed that as incitement to genocide is an ‘inchoate crime’, genocide need not actually occur since the inciting utterance itself constitutes the criminal conduct.

²⁰³ *Ibid.*, §§698, 701. See also the Partly Dissenting Opinion of Judge Shahabuddeen, which provides that the ‘incitement must call for immediate action, but it certainly is not the case that the prosecution has to show that genocide in fact followed immediately after the message or at all’: §65.

²⁰⁴ *Ibid.*, §692 (defining such speech as ‘inciting discrimination or violence’). See also §693 (noting that as a result ‘the jurisprudence on incitement to hatred, discrimination and violence is not directly applicable in determining what constitutes direct incitement to commit genocide’).

international law that criminalizes hate speech, as the ‘number and extent of the reservations’ to article 4 of CERD and article 20 of the ICCPR reveal ‘that profound disagreement persists in the international community as to whether mere hate speech [as opposed to incitement to genocide] is or should be prohibited.’²⁰⁵

In *Nahimana* and subsequent decisions, the ICTR confirmed that the crime of incitement to genocide does not incorporate a requirement of imminent or likely harm and instead focuses on the intent of the speaker, thereby diverging from the mandatory restrictions in article 20(2) of the ICCPR and article 4 of CERD on this point. The *travaux préparatoires* illustrate that this was a point of tension in the drafting history of the Convention, with the United States putting forward alternative text which criminalized incitement to genocide only when ‘such incitement takes place under circumstances which may reasonably result in the commission of acts of genocide.’²⁰⁶ In *Nahimana*, the Appeals Chamber noted that although stringent *mens rea* requirements applied, ‘direct and public incitement to commit genocide . . . is itself a crime, and it is not necessary to demonstrate that it in fact substantially contributed to the commission of acts of genocide.’²⁰⁷ This was reiterated in *Nzabonimana*, which concluded that ‘the *actus reus* of direct and public incitement to genocide is satisfied when a person directly and publicly incites the commission of genocide, *irrespective of whether his or her acts were likely to cause the crime of genocide*.’²⁰⁸

However, some scholars have observed that ‘a number of ICTR judgments felt it necessary to establish a direct causal link between speeches or broadcasts and actual genocide or other material crimes in the “findings of fact” sections’, such as the *Nahimana* case, where the ruling ‘asserts no less than sixteen times that speech acts directly caused genocidal killings.’²⁰⁹ They argue that the ICTR seems to have thereby introduced ‘a new criterion of contemporaneousness, in which genocidal speech acts had to be uttered very near or simultaneous with the onset of actual genocide’ and made causation a requirement in its incitement jurisprudence after all.²¹⁰

1.4. Mandatory restrictions at the regional level

The American Convention is the only regional human rights instrument containing a mandatory requirement to penalize specific speech. Article 13(5) of the American Convention provides that ‘[a]ny propaganda for war and any advocacy of national,

²⁰⁵ Ibid., Partly Dissenting Opinion of Judge Meron, §§5–8. Judge Meron held that the Appeals Chamber impermissibly predicated convictions for crimes against humanity by persecution on the basis of hate speech, rather than the incitement to commit genocide convictions.

²⁰⁶ Prevention and Punishment in the draft convention prepared by the Secretariat, Lake Success, N.Y. (1948) UN Doc. E/623 and Ad Hoc Committee on Genocide, Basic Principles of a Convention on Genocide Proposed by the Delegation of the Union of Soviet Socialist Republics on 5 April 1948 (1948) UN Doc. E/AC.25/7, cited in W. Schabas, *Genocide in International Law: The Crime of Crimes* (1st edn, CUP 2000), 267. The USSR argued for broader versions of the provision that criminalized ‘inciting racial, national or religious enmity or hatred’: 271.

²⁰⁷ ICTR, *Prosecutor v. Nahimana et. al.* (ICTR-99-52-A), Judgement, 28 November 2007, §678.

²⁰⁸ ICTR, *Nzabonimana v. Prosecutor* (ICTR-98-44D-A), Judgement, 29 September 2014, §234 (emphasis added).

²⁰⁹ See Wilson (n 197) 34–37.

²¹⁰ Ibid. See also S. Benesch, ‘Vile Crime in Inalienable Right: Defining Incitement to Genocide’ (2008) 48(3) *Virginia Journal of International Law* 485, 497 (‘courts have indirectly relied on causation, in the absence of such a test, even where such reliance may have stretched the evidence thin’).

racial, or religious hatred that constitute incitements to lawless violence or to any other similar 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.²¹¹

The mandatory hate speech restriction in the Inter-American system is in some ways similar to the one contained in the ICCPR and CERD. It is subject to the three-part test set forth in the jurisprudence of the Inter-American Court of Human Rights which, like the ICCPR article 19(3) test, requires an assessment of legality, legitimacy and necessity.²¹² And article 13(5) appears to incorporate similar intent and imminence standards as the ICCPR and CERD provisions, requiring that states must demonstrate ‘actual, truthful, objective and strong proof that the person was not simply issuing an opinion (even if that opinion was hard, unfair or disturbing), but that the person had the clear intention of committing a crime and the actual, real and effective possibility of achieving this objective’ before their speech can be criminalized.²¹³

In one sense the Inter-American provision can capture more speech than the other treaties, because the list of target groups is wider: according to the Inter-American Commission, article 13(5) covers characteristics other than nationality, race and religion and extends to speech that is hateful on the basis of ‘sexual orientation, gender identity, and bodily diversity.’²¹⁴

But the harm threshold provided by article 13(5)—‘lawless violence or other similar action’ is much higher—and therefore more speech-protective—than the equivalent treaties.²¹⁵ Though ‘similar action’ is not defined, the drafting history of the treaty suggests that the harm should be comparable in severity to violence.²¹⁶ And in relation to

²¹¹ ACHR Art. 13(5). However, as the OAS Special Rapporteur for Freedom of Expression has previously noted, there is a ‘discrepancy’ between the English and Spanish language versions of the text of article 13. The Spanish language version of article 13(5) – ‘*estará prohibida por la ley*’ – suggests that speech must be ‘prohibited’ rather than ‘punishable’. But the Inter-American Court is yet to comprehensively address the impact of this discrepancy – and whether it calls into question the mandatory nature of criminal restrictions under article 13(5) – in its jurisprudence: OAS Special Rapporteur for Freedom of Expression, *Hate Speech and the American Convention* (2004), §§37–38.

²¹² See, e.g., IACtHR, *Herrera-Ulloa v. Costa Rica* (Series C, no. 107), 2 July 2004, §123. See also IACmHR, *Annual Report of the Inter-American Commission on Human Rights* (2009) OEA/Ser.L/V/II., Doc. 51, 247, 252 (stating that “‘necessary” is not synonymous with “useful”, “reasonable” or “convenient”” but requires states to use the least intrusive means to achieve their legitimate public interest purpose and that ‘proportionate’ means assessing ‘whether the sacrifice of freedom of expression ... is excessive in relation to the advantages obtained through such measure’). Like the HRC, the IACmHR has clarified that the burden is on ‘the authority imposing limitations to prove that these conditions have been met’: 247.

²¹³ IACmHR, *Inter-American Legal Framework Regarding the Right to Freedom of Expression* (2009) OEA/Ser.L/V/II, CIDH/RELE/INF. 2/09, §59.

²¹⁴ IACmHR, *Violence against Lesbian, Gay, Bisexual, Trans and Intersex Persons in the Americas* (2015) OAS/Ser.L/V/II. Doc.36/15 Rev.2, §13, referring to ‘the [IACmHR] and its Special Rapporteurship affirm that article 13(5) includes hate speech that incites lawless violence against a group’ on those additional grounds. See s. VI. (Recommendations).

²¹⁵ See, e.g., OAS Special Rapporteur for Freedom of Expression, *Hate Speech and the American Convention* (2004), §46 (observing that article 13(5) ‘suggest[s] that violence is a requirement for any restrictions’ but that the ICCPR and European Convention ‘do not have such a narrowly drawn requirement’).

²¹⁶ Inter-American specialized conference on human rights, San José, Costa Rica, Minutes of the second plenary session (22 November 1969) Doc. 86, US Statement, 444. The original provision provided for violence as well as ‘discrimination’ and ‘hostility’, but the US delegate recommended that this wording be replaced with ‘similar action’ to ensure consistency with First Amendment jurisprudence. In addition the Spanish version of article 13(5) provides for ‘similar *illegal action*’ (‘acción ilegal similar’).

the causal link between speech and harm under article 13(5), it is not permissible to ‘invoke as a reason to limit freedom of expression mere conjectures about eventual effects on the public order, or hypothetical circumstances derived from subjective interpretations by authorities of facts that do not clearly present a present, certain, objective and imminent risk of violence.’²¹⁷ Instead, a speaker must have a ‘real, present and effective possibility of achieving his or her objectives.’²¹⁸

1.5. Summary

In summary, the requirements in international and regional treaties that states impose mandatory penalties for hate speech, although not identical, share key common elements. Every mandatory penalty must comply with the permissible limitations to freedom of expression. Each infers an intent requirement, with article 3 of the Genocide Convention requiring the highest intent standard linked to the act of genocide. And each treaty requires speech to rise to the level of ‘incitement’, thereby incorporating an element of imminent risk of harm, despite article 4 of CERD on its face allowing a lower threshold of ‘dissemination of ideas’ only.²¹⁹

The principal variation is in the harms that are recognized as legitimate bases for criminal penalties for speech. The recognized harms include discrimination, hostility or violence in the case of article 20 ICCPR; hatred, discrimination, violence and contempt under article 4 CERD; and ‘lawless violence’ or ‘other similar action’ in the Inter-American provision. The American Convention sets the highest bar, though there has been minimal authoritative guidance defining the contours of these terms in the jurisprudence of the Inter-American Court to date.

The targeted group each harm relates to also diverges, with the CERD only applying to ‘racial’ origin,²²⁰ the ICCPR also including ‘religious’ groups and the American Convention—as interpreted by the Inter-American Commission—extending further to groups defined by gender and sexual orientation.²²¹ Finally, the penalties differ, with article 4 of CERD, article 3 of the Genocide Convention and article 13(5) of the American Convention²²² requiring criminalization, not just civil sanctions, for speech.

²¹⁷ IACmHR, *Democracy and Human Rights in Venezuela* (2009) OEA/Ser.L/V/II., Doc. 54, §379; IACmHR, *Annual Report of the Inter-American Commission on Human Rights* (2009) OEA/Ser.L/V/II., Doc. 51, 169.

²¹⁸ IACmHR, *Democracy and Human Rights in Venezuela* (2009) OEA/Ser.L/V/II., Doc. 54, §360; IACmHR, *Annual Report of the Inter-American Commission on Human Rights* (2009) OEA/Ser.L/V/II., Doc. 51, 174.

²¹⁹ See s. III.1.2. (CERD Article 4). Although article 3 of the Genocide Convention requires an element of ‘incites’, this has been held not to equate to a requirement of imminent risk of harm: See s. III.1.3. (Genocide Convention Article 3).

²²⁰ The CERD Committee has, however, interpreted this to extend to ‘colour, descent or national or ethnic origin’: CERD Committee, General Recommendation No. 35 (2013), §§6, 13. See also CERD Art. 1.

²²¹ See s. III.1.2. (CERD Article 4); see also CERD Committee, General Recommendation No. 35 (2013), §13; ICCPR Art. 20; IACmHR, *Violence against Lesbian, Gay, Bisexual, Trans and Intersex Persons in the Americas* (2015) OAS/Ser.L/V/II. Doc.36/15 Rev.2, §13.

²²² The IACtHR has rejected the use of criminal law to punish expressions in cases involving the public interest, and has favoured the use of the right of reply and civil damages to deal with abuses of freedom of expression. But the IACtHR has not ruled on cases of hate speech specifically. See IACtHR, *Álvarez Ramos v. Venezuela* (Series C, no. 380), 30 August 2019, §§120–122; see s. III.2.5. (Penalties).

2. Discretionary Restrictions on Hate Speech

In addition to treaty provisions governing when states *must* punish hate speech,²²³ international and regional human rights law provides the legal framework for when states *may* do so. Under article 19(3) of the ICCPR, hate speech can be ‘subject to certain restrictions’ only as ‘provided by law’ and as ‘necessary’ for the protection of certain legitimate aims including the right to reputation, national security and public health. This tripartite test is often referred to as one of legality, legitimacy and necessity.²²⁴ An element of the requirement of necessity is proportionality: any restrictive measure ‘must be the least intrusive instrument amongst those which might achieve their protective function’ and must be ‘proportionate to the interest to be protected.’²²⁵

Article 10 of the European Convention sets out a tripartite test similar to article 19(3) of the ICCPR, but also provides a number of legitimate aims not explicitly recognized in article 19(3) of the ICCPR. These are: ‘territorial integrity or public safety’, the ‘prevention of disorder or crime’, ‘preventing the disclosure of information received in confidence’ and ‘maintaining the authority and impartiality of the judiciary’. However, the Human Rights Committee has generally construed the limitations outlined in article 19(3) as encompassing the aims set out in article 10(2), for instance by way of a broad reading of the ‘public order’ limitation.²²⁶

The list of legitimate aims articulated in article 13(2) of the American Convention mirrors article 19(3) of the ICCPR and the Inter-American Court applies the three-part test of legality, legitimacy and necessity consistently with other regional bodies.²²⁷ The American Convention is, however, more protective of speech than its equivalents in its explicit prohibition of prior restraints to speech in almost all circumstances,²²⁸ as well as indirect government restrictions and ‘private controls.’²²⁹ The Inter-American Court has indeed explicitly stated that a textual comparison of article 13 with article 10 of the European Convention and article 19 of the ICCPR ‘indicates clearly that the guarantees contained in the American Convention regarding freedom of expression were

²²³ See s. III.1. (Mandatory Restrictions on Hate Speech).

²²⁴ UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2016) UN Doc. A/71/373, §57(a).

²²⁵ HRC, General Comment No. 34 (2011), §34.

²²⁶ See ch. 1 (Introduction), s. II.1.1.1. (Abuses by the state). See also HRC, *Lovell v. Australia* (Comm. no. 920/2000), 24 March 2004, §9.4; HRC, General Comment No. 34 (2011), §31; HRC, *Agazade and Jafarov v. Azerbaijan* (Comm. no. 2205/2012), 27 October 2016, §7.4; HRC, *Aleksandrov v. Belarus* (Comm. no. 1933/2010), 24 July 2014, §7.4.

²²⁷ See, e.g., IACtHR, *Herrera-Ulloa v. Costa Rica* (Series C, no. 107), 2 July 2004, §120.

²²⁸ ACHR Art. 13(2) (‘The exercise of the right [to freedom of thought and expression] shall not be subject to prior censorship’), Art. 13(4) (‘Notwithstanding the provisions of paragraph 2 above, public entertainments may be subject by law to prior censorship for the sole purpose of regulating access to them for the moral protection of childhood and adolescence’). See also IACtHR, *Compulsory Membership in an Association Prescribed by Law for the Practice of Journalism (Arts. 13 and 29 American Convention on Human Rights)*, Advisory Opinion OC-5/85 (Series A, no. 5), 13 November 1985, §38 (‘Article 13 (2) . . . stipulates, in the first place, that prior censorship is always incompatible with the full enjoyment of the rights listed in Article 13, but for the exception provided for in subparagraph 4 dealing with public entertainments, even if the alleged purpose of such prior censorship is to prevent abuses of freedom of expression’).

²²⁹ ACHR Art. 13(3).

designed to be more generous and to reduce to a bare minimum restrictions impeding the free circulation of ideas.²³⁰ While this speech-protective approach has not yet been applied to hate speech, it is indicative of the approach likely to be taken by the Inter-American Court.

Article 9(2) of the African Charter is unique in its lack of detail: the provision states only that '[e]very individual shall have the right to express and disseminate his opinions within the law'. However, the African Court has held that the term 'within the law' means that restrictions on speech must be 'prescribed by law, serve a legitimate purpose and [be] necessary and proportional as may be expected in a democratic society'.²³¹ The Court has also held that article 9(2) 'must be interpreted in the light of international human rights standards',²³² and its decisions related to the right to free speech draw on article 19 of the ICCPR to determine whether restrictions to speech violate the African Charter.²³³

The European Court of Human Rights has, for its part, emerged as an outlier in the extent to which it allows the penalization and even criminalization of hate speech, including Holocaust denial, religious insults and speech that does not incite violence. This is somewhat ironic given that other systems require mandatory prohibitions on such speech under certain circumstances, while the European Convention does not.²³⁴

There are two key divergences between the European Court and its counterparts that result in the Court being a less speech-protective human rights body. First, unlike the Human Rights Committee,²³⁵ the European Court considers that states 'have a certain margin of appreciation' in assessing whether a 'pressing social need' makes a speech restriction 'necessary' to protect the objectives articulated in article 10(2) of the European Convention.²³⁶ Although the European Court applies a narrow margin where speech relates to a matter of public concern,²³⁷ it has also stated it affords governments a particularly wide margin of appreciation when expression is 'liable to offend intimate personal convictions within the sphere of morals or, especially, religion',²³⁸ and 'where such remarks incite violence against an individual, a public official or a sector

²³⁰ IACtHR, *Compulsory Membership in an Association Prescribed by Law for the Practice of Journalism (Arts. 13 and 29 American Convention on Human Rights)*, Advisory Opinion OC-5/85 (Series A, no. 5), 13 November 1985, §50.

²³¹ ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017, §133.

²³² *Ibid.*, §136. Article 60 of the ACHR provides that the African Commission 'must draw inspiration from international law on human and peoples' rights'. Although article 60 refers to the African Commission only, as the ACtHPR was established in 1998 after the Charter was adopted, the ACtHPR has since interpreted article 60 to grant the Court the same powers: see, e.g., ACtHPR, *Omary and others v. Tanzania* (App. no. 001/2012), 28 March 2014, §73 ('Article 60 of the Charter empowers the Court to "draw inspiration from international law on human and peoples' rights"').

²³³ See, e.g., ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017, §163; ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §164.

²³⁴ See s. III.1. (Mandatory Restrictions on Hate Speech).

²³⁵ HRC, General Comment No. 34 (2011), §36.

²³⁶ See, e.g., ECtHR (GC), *Zana v. Turkey* (App. no. 18954/91), 25 November 1997, §51. However, this margin is more 'limited' when freedom of the press is at stake: see s. III.2.4.4. ('Responsible journalism').

²³⁷ ECtHR, *Otegi Mondragon v. Spain* (App. no. 2034/07), 15 March 2011, §51; ECtHR, *Tagiyev v. Azerbaijan* (App. no. 13274/08), 5 December 2019, §36.

²³⁸ ECtHR, *Murphy v. Ireland* (App. no. 44179/98), 10 July 2003, §67. See also ECtHR, *Wingrove v. United Kingdom* (App. no. 17419/90), 25 November 1996, §58.

of the population.²³⁹ By contrast, the Human Rights Committee has explicitly rejected such an approach, observing that the scope of freedom of expression ‘is not to be assessed by reference to a “margin of appreciation”’.²⁴⁰

The second critical difference between the European Court and its counterparts is the use of article 17 of the European Convention in cases involving speech. Article 17 provides that ‘[n]othing in th[e] Convention may be interpreted as implying ... any right to engage in any activity ... aimed at the destruction of any of the rights and freedoms’ set out in the treaty, and the Court regularly dismisses claims related to hate speech as being non-admissible on this basis. This means that in some cases speech is not analysed using the balancing exercise provided by article 10(2).²⁴¹ The Court has construed article 17 as preventing ‘individuals or groups with totalitarian aims from exploiting ... the principles enunciated in the Convention.’²⁴² And the Court has, on this basis, refused to consider claims involving ‘speech [that] is incompatible with the values’ of the Convention or ‘contrary to’ its ‘text and spirit.’²⁴³ Although the Court has stated that article 17 sets a ‘high threshold’²⁴⁴ and should only be applied ‘on an exceptional basis and in extreme cases,’²⁴⁵ it has in effect operated as a form of guillotine provision which overrides the balancing exercise applicable to other types of speech.²⁴⁶ And, ultimately, the 46 member states of the European Convention are also signatories to the ICCPR, and to the extent that their restrictions to speech are incompatible with article 19(3), they will be in breach of their international legal obligations.²⁴⁷

2.1. Legality

According to the Human Rights Committee, the requirement in article 19(3) of the ICCPR that a restriction on speech be ‘provided by law’ means that it must be public and ‘formulated with sufficient precision to enable an individual to regulate his or her conduct.’²⁴⁸

²³⁹ ECtHR, *Savva Terentyev v. Russia* (App. no. 10692/09), 28 August 2018, §65.

²⁴⁰ HRC, General Comment No. 34 (2011), §36. The IACtHR and IACmHR have previously referred to states having a margin of appreciation, but have generally ‘applied this doctrine with great caution and in many fewer instances than their European counterparts’, and in relation to freedom of expression have ‘been very hesitant to grant deference to States, preferring to examine closely the scope and effect of restrictive domestic measures’: see B. Duhaime, ‘Subsidiarity in the Americas: What Room Is There for Deference in the Inter-American System?’, in L. Gruszczynski & W. Werner (eds), *Deference in International Courts and Tribunals: Standard of Review and Margin of Appreciation* (OUP 2014), 301, 307.

²⁴¹ See ECtHR, ‘Factsheet—Hate Speech’ (September 2023).

²⁴² ECtHR, *Hizb Ut-Tahrir v. Germany* (App. no. 31098/08), 12 June 2012, §72, citing ECtHR (GC), *Paksas v. Lithuania* (App. no. 34932/04), 6 January 2011, §§87–88.

²⁴³ ECtHR, *Ivanov v. Russia* (App. no. 35222/04), 20 February 2007, §1; ECtHR, *Roj TV A/S v. Denmark* (App. no. 24683/14), 24 May 2018, §30.

²⁴⁴ ECtHR, *Lilliendahl v. Iceland* (App. no. 29297/18), 12 May 2020, §§25–26.

²⁴⁵ *Ibid.*, §25. See also ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §114.

²⁴⁶ ECtHR, *Pastörs v. Germany* (App. no. 55225/14), 3 October 2019, §36; UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §26.

²⁴⁷ Russia having ceased to be a party to the European Convention on Human Rights on 16 September 2022. Council of Europe, ‘Russia ceases to be party to the European Convention on Human Rights’ (16 September 2022). UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §26. See ch 1 (Introduction), s. II.3. (Jurisprudence).

²⁴⁸ HRC, General Comment No. 34 (2011), §25. Norms enshrined in traditional, religious or other customary laws are not sufficient: HRC, General Comment No. 34 (2011), §24.

The Committee has urged legislators to ‘avoid’ the ‘vague terminology and overly broad restrictions’ that exist in many states.²⁴⁹ It has, for instance, concluded that Russia should ‘clarify the vague, broad and open-ended definition of key terms’ in laws criminalizing ‘blasphemy.’²⁵⁰ And it found that the offence of ‘promoting propaganda of homosexuality’ was ‘highly ambiguous as to the actions being prohibited and therefore [did] not satisfy the requirement of lawfulness under article 19(3).’²⁵¹ It has also requested that Kazakhstan ‘[c]larify the vague and broad definition of key terms’ in its laws restricting speech, including the offence of incitement to ‘social, national, clan, class or religious discord’ to comply with article 19 of the ICCPR.²⁵²

Like the Human Rights Committee, the UN’s Working Group on Arbitrary Detention has insisted on the need for precise laws—including in a case where a woman in Iran was convicted of ‘spreading propaganda’ and ‘insulting the sanctity of Islam’ for posts supporting Iranian dissidents and an entry in her private journal about a woman who burned a copy of the Qur’an.²⁵³ The Working Group considered that the charges were ‘so vague and overly broad’ that she could not have foreseen a criminal penalty on this basis and that, as a result, her deprivation of liberty violated ‘the principle of legality.’²⁵⁴

Similarly, the European Court has interpreted the legality (or ‘prescribed by law’) test in article 10(2) of the European Convention as requiring a law ‘formulated with sufficient precision to enable the person concerned to regulate his or her conduct’ and to ‘foresee, to a degree that was reasonable in the circumstances, the consequences that a given action could entail.’²⁵⁵ However, the Court has applied the legality test loosely, in some instances finding vague incitement laws that punish individuals for ‘express[ing] contempt’ or ‘mock[ing]’ persons for certain characteristics as having satisfied the legality test.²⁵⁶ The European Court has also at times bypassed the legality test entirely by finding that vagueness concerns would be dealt with as part of its analysis of whether the interference with speech was necessary and proportionate.²⁵⁷

The Inter-American Court applies a stricter standard of legality than the European Court—more akin to the Human Rights Committee’s jurisprudence—particularly when a state is seeking to limit speech through criminal sanctions. In these circumstances, ‘it is necessary to use strict and unequivocal terms’ to describe speech that can

²⁴⁹ HRC, *Concluding Observations, Cambodia* (2015) UN Doc. CCPR/C/KHM/CO/2, §21. See s. II.1.1. (Type of speech).

²⁵⁰ HRC, *Concluding Observations, Russian Federation* (2015) UN Doc. CCPR/C/RUS/CO/7, §§19–20 (Terrorism and public order offences are most commonly criticized by the Committee as impermissibly vague). See also ch. 6 (Terrorism Laws), s. III.4.2. (Truth and opinion).

²⁵¹ HRC, *Nepomnyashchiy v. Russian Federation* (Comm. no. 2318/2013), 17 July 2018, §7.7.

²⁵² HRC, *Concluding Observations, Kazakhstan* (2016) UN Doc. CCPR/C/KAZ/CO/2, §§49–50.

²⁵³ WGAD, *Iraee v. Islamic Republic of Iran* (Opinion no. 33/2019), 12 August 2019, §51.

²⁵⁴ *Ibid.*, §§51–53 (also finding a violation under article 19 of the UDHR).

²⁵⁵ ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §131 (citing ECtHR, *Sunday Times v. United Kingdom* (App. no. 6538/74), §§48–49).

²⁵⁶ See, e.g., ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012, §§24, 49; ECtHR, *Lilliendahl v. Iceland* (App. no. 29297/18), 12 May 2020, §§20, 42.

²⁵⁷ See, e.g., ECtHR, *Otegi Mondragon v. Spain* (App. no. 2034/07), 15 March 2011, §§35, 45–46. See also ECtHR, *Savva Terentyev v. Russia* (App. No. 10692/09), 28 August 2018, §§42, 58.

be incriminated, ‘setting its elements, and defining the behaviors that are not punishable or the illicit behaviors that can be punishable with non-criminal measures’.²⁵⁸

Unlike the other treaties, article 9 of the African Charter does not set out specific exceptions to the right to ‘express and disseminate his opinions’. However, article 9(2) requires that this right be exercised ‘within the law’, and the African Court and Commission have held that this phrase ‘must be interpreted in reference to international norms which can provide grounds of limitation on freedom of expression’, thereby incorporating the principles of legality, legitimacy and necessity into the African Charter.²⁵⁹ Adopting the Human Rights Committee’s reasoning as to ‘legality’, the African Court has held that laws must be drafted with ‘sufficient clarity’ to enable an individual to adapt his or her conduct.²⁶⁰ But in a decision regarding Rwandan laws criminalizing the ‘minimization’ or ‘downplaying’ of genocide,²⁶¹ the African Court recognized that although these laws were ‘couched in broad and general terms’, ‘the nature of the offences ... is admittedly difficult to specify with precision’ and that, ‘considering the margin of appreciation that [the state] enjoys in defining and prohibiting some criminal acts in its domestic legislation, the Court is of the view that the impugned laws provide adequate notice for individuals to foresee and adapt their behaviour to the rules’.²⁶²

2.2. Legitimacy

Under UN human rights standards, any restriction on expression must fall into one of ‘two limitative areas of restrictions’: it must be aimed at (i) respecting the rights or reputation of others, or (ii) protecting national security, public order, public health or morality.²⁶³ The European and Inter-American regional human rights instruments include similar requirements, and the African Charter has been construed as encompassing these legitimate aims as well.²⁶⁴

2.2.1. Human Rights Committee

The Human Rights Committee has emphasized that the legitimate aims enumerated in article 19(3) of the ICCPR are exhaustive and cannot be used as a pretext for other objectives.²⁶⁵ And the burden of proof lies on the government to demonstrate that any

²⁵⁸ IACtHR, *Usón Ramírez v. Venezuela* (Series C, no. 207), 20 November 2009, §55: see ch. 5 (Espionage and Official Secrets Laws), s. III.2. (Legality).

²⁵⁹ ACtHPR, *Konaté v. Burkina Faso* (App. No. 004/2013), 5 December 2014, §129, citing ACmHPR, *Malawi African Association and Others v. Mauritania* (Comm. No. 54/91 and others), 11 May 2000, §102. See also ACtHPR, *Umuhoza v. Rwanda* (App. No. 003/2014), 24 November 2017, §133. See ch. 1 (Introduction), s. II.1.1.1. (Abuses by the state).

²⁶⁰ ACtHPR, *Konaté v. Burkina Faso* (App. No. 004/2013), 5 December 2014, §131.

²⁶¹ ACtHPR, *Umuhoza v. Rwanda* (App. No. 003/2014), 24 November 2017, §137, 155.

²⁶² *Ibid.*, §§137–138. The Court ultimately held that a 15-year term of imprisonment under these laws for a political speech at a Genocide Memorial was not necessary in a democratic society, as the speech was offensive but could not reasonably be considered capable of ‘inciting strife’, creating ‘divisions among people’ or ‘threatening the security of the State’: *ibid.*, §§161–162. See ss II.1.1.1. (Type of speech) (Case study: Denial laws) and III.2.4.2. (Opinion).

²⁶³ HRC, General Comment No. 34 (2011), §21.

²⁶⁴ See s. III.1. (Mandatory Restrictions on Hate Speech).

²⁶⁵ HRC, General Comment No. 34 (2011), §§22–30.

restriction to speech, including hate speech, satisfies the tests of legality, legitimacy and necessity.²⁶⁶

Respondent states in the small number of hate speech cases that have come before the Human Rights Committee have argued that restrictions on speech were justified to show respect for the rights of others.²⁶⁷ The Committee has held in these cases that the term ‘rights’ ‘includes human rights as recognized in the Covenant and ... international human rights law’, and that ‘others’ relates to ‘other persons individually or as members of the community’.²⁶⁸ In a case in which a school teacher was transferred to a non-teaching position due to publications calling on Christians to ‘question the validity of Jewish beliefs and teachings’ and ‘hold those of the Jewish faith and ancestry in contempt’, the Committee held that the penalty was justified for the purpose of ‘protecting the “rights or reputations” of persons of Jewish faith’. This included the right to have ‘an education in the public school system free from bias, prejudice and intolerance’.²⁶⁹

2.2.2. *European Court of Human Rights*

Consistently with the Human Rights Committee’s jurisprudence, hate speech is generally justified before the European Court under article 10(2) on the basis of ‘protection of the rights or reputation of others’.²⁷⁰ This notion has been construed broadly. For example, when a Russian author wrote a demeaning article about non-Russians but did not identify a particular ethnic group, the Court accepted that his conviction for inciting hatred and enmity was aimed at protecting the ‘“rights of others”, specifically the dignity of people of a non-Russian ethnicity’.²⁷¹ When hate speech is ‘capable of inciting to further violence’, restrictions will more commonly be justified by reference to the maintenance of national security, territorial integrity or public safety, or for the prevention of disorder or crime.²⁷²

²⁶⁶ *Ibid.*, §27.

²⁶⁷ See HRC, *Ross v. Canada* (Comm. No. 736/1997), 18 October 2000, §§6.11–6.12. Cf. HRC, *A.K. and A.R. v. Uzbekistan* (Comm. No. 1233/2003), 31 March 2009, §7.2 (where a conviction for incitement of ethnic, racial, or religious hatred was held to be justified on the basis of a ‘perceived threat to national security (violent overthrow of the constitutional order) and to the rights of others’). See ch. 1 (Introduction), s. II.3.1.1. (Justifications for penalizing speech).

²⁶⁸ HRC, General Comment No. 34 (2011), §28. See also HRC, *Ross v. Canada* (Comm. No. 736/1997), 18 October 2000, §11.5; HRC, *Faurisson v. France* (Comm. No. 550/1993), 8 November 1996, §9.6.

²⁶⁹ HRC, *Ross v. Canada* (Comm. No. 736/1997), 18 October 2000, §§11.5–11.6. The Committee ultimately held that the restriction ‘did not go any further than that which was necessary to achieve its protective functions’, and therefore that there had been no violation of article 19. See also HRC, *Faurisson v. France* (Comm. No. 550/1993), 8 November 1996 (although this ruling was essentially overturned in HRC, General Comment No. 34 (2011): see ss II.1.1. (Type of speech) (Case Study: Denial laws) and III.2.4.2. (Opinion)).

²⁷⁰ See, e.g., ECtHR, *Lilliendahl v. Iceland* (App. No. 29297/18) 12 May 2020, §43; ECtHR, *Balsytė-Lideikienė v. Lithuania* (App. No. 72596/01), 4 November 2008, §73; ECtHR (GC), *Jersild v. Denmark* (App. No. 15890/89), 23 September 1994, §27 (where this legitimate aim was ‘uncontested’). There are also a number of cases where the European Court has found a restriction justified on the basis of the rights of others as well as additional legitimate aims: see, e.g., ECtHR, *Féret v. Belgium* (App. No. 15615/07), 16 July 2009, §59 (holding that a criminal conviction for leaflets presenting non-European immigrants as criminally-minded protected the rights of others and the prevention of disorder).

²⁷¹ ECtHR, *Atamanchuk v. Russia* (App. No. 4493/11), 11 February 2020, §42.

²⁷² See e.g., ECtHR (GC), *Sürek v. Turkey (No. 1)* (App. No. 26682/95), 8 July 1999, §§51–52, 62. See also ch. 6 (Terrorism Laws), s. III.2. (Legitimacy).

In cases where the European Court has assessed Holocaust-denial laws, it has held that such laws pursue the legitimate aim of protecting the rights and reputation of others, as well as in some instances the ‘prevention of disorder.’²⁷³ For example, when a bishop denied the existence of the Holocaust in a TV interview and was charged under Germany’s incitement-to-hatred law, the Court considered that this law ‘pursued the legitimate aim of preventing a disturbance of the public peace in Germany and thus the prevention of disorder and crime.’²⁷⁴ But the Court reached a different conclusion in a case in which a politician declared at a public event that the Armenian genocide was a ‘lie.’ Although the Court accepted that his criminal conviction for stirring up racial discrimination was justified on the basis of the rights of others—namely the ‘dignity of present-day Armenians’ who have constructed an ‘identity around the perception that their community has been the victim of a genocide’—the Court did not accept that such speech should be restricted on the basis of preventing disorder.²⁷⁵ Instead, it construed ‘the prevention of disorder’ as an objective that narrowly encompasses ‘riots or other forms of public disturbance’ and held that there was no evidence ‘this kind of statement could risk unleashing serious tensions and giving rise to clashes’ given the timing and location of the speech.²⁷⁶

Certain types of hate speech have also been found to fall within article 17—meaning that they are by definition unable to meet the ‘legitimate aim’ test and the case is deemed inadmissible. Speech has been considered ineligible for protection by the Court under this article on the basis that it contravenes values guaranteed by the Convention such as ‘justice and peace.’²⁷⁷ For example, the Court held that anti-Jewish speech constituted a ‘general and vehement attack on one ethnic group’ in a manner that violates the Convention’s ‘underlying values’ of ‘tolerance, social peace and non-discrimination.’²⁷⁸ Displaying a poster showing the Twin Towers in flames and stating ‘Islam out of Britain—Protect British People’ in the speaker’s home was considered contrary to the same values and therefore ineligible for protection as it constituted an attack ‘against a religious group, linking the group as a whole with a grave act of terrorism.’²⁷⁹ And calls for violence such as speech on television that ‘included incitement to violence and support for terrorist activity’ and YouTube videos ‘call[ing] on viewers to overpower and fight non-Muslims’ have been considered contrary to the ‘values’ of the Convention including tolerance, social peace and non-discrimination.²⁸⁰

²⁷³ See, e.g., ECtHR, *Pastörs v. Germany* (App. No. 55225/14), 3 October 2019, §41; ECtHR, *Williamson v. Germany* (App. No. 64496/17), 8 January 2019, §24. See s. III.2.4.2. (Opinion).

²⁷⁴ ECtHR, *Williamson v. Germany* (App. No. 64496/17), 8 January 2019, §24.

²⁷⁵ ECtHR (GC), *Perinçek v. Switzerland* (App. No. 27510/08), 15 October 2015, §§154, 156.

²⁷⁶ *Ibid.*, §§146–154. The Court ultimately held that a violation of article 10 had taken place: see s. III.2.4.2. (Opinion).

²⁷⁷ See s. III.2.4.2. (Opinion). ECtHR, *Garaudy v. France* (App. No. 65831/01), 24 June 2003, §1. Cf. ECtHR (GC), *Perinçek v. Switzerland* (App. No. 27510/08), 15 October 2015.

²⁷⁸ ECtHR, *Ivanov v. Russia* (App. No. 35222/04), 20 February 2007, §1.

²⁷⁹ ECtHR, *Norwood v. United Kingdom* (App. no. 23131/03), 16 November 2004.

²⁸⁰ ECtHR, *Belkacem v. Belgium* (App. No. 34367/14), 27 June 2017, §§5–7; ECtHR, *Roj TV A/S v. Denmark* (App. No. 24683/14), 17 April 2018, §§47–48. See also ECtHR, *Hizb Ut-Tahrir v. Germany* (App. No. 31098/08), 12 June 2012, §72. Cf. ECtHR, *Ifandiiev v. Bulgaria* (App. No. 14904/11), 18 April 2019, §24 (declining to apply article 17 unless ‘it is immediately clear’ that the impugned speech sought to ‘deflect [article 10] from its real purpose by employing the right to freedom of expression for ends clearly contrary to the values of the Convention’); ECtHR, *Church of Scientology Moscow and Others v. Russia* (App. No. 37508/1 & 2 others), 14 December 2021, §§35, 59.

2.2.3. Inter-American Court and Commission

The Inter-American Court has not directly opined on restrictions to hate speech, but the Court has considered that ‘respect for the rights or reputations of others’ is a legitimate aim that justified certain restrictions to speech.²⁸¹ And the Court has recognized that ‘the protection of a person’s honor and reputation is a legitimate end,’²⁸² and that ‘it is necessary that the State weigh up the right to freedom of expression of the communicator and the right to honor of the person affected.’²⁸³

2.2.4. African Charter

The African Court has limited jurisprudence on restrictions to hate speech²⁸⁴ but the phrase ‘within the law’ in article 9 of the African Charter has been interpreted to require that ‘restrictions are prescribed by law, serve a legitimate purpose and are necessary and proportional as may be expected in a democratic society,’ thereby bringing the provision into line with international and regional equivalents.²⁸⁵

The African Court and Commission and the Economic Community of West African States Community Court of Justice have indeed specifically held that the aims listed in article 19(3) of the ICCPR are ‘legitimate’ grounds to restrict speech and these aims are stated to be exhaustive.²⁸⁶ But it is not clear whether additional aims are also recognized. Article 27(2) of the African Charter provides that Charter rights generally ‘shall be exercised with due regard to the rights of others, *collective security, morality and common interest*’ and it is unclear whether the aims in article 19(3) of the ICCPR and article 27(2) of the Charter are considered cumulative or not. This uncertainty is

61 (declining to apply article 17 in a case involving a ban on speech promoting Scientology because there was no evidence ‘that the impugned texts . . . used abusive terms’ about persons who didn’t believe in Scientology); ECtHR, *Ibragimov v. Russia* (App. nos. 1413/08 & 28621/11), 28 August 2018, §123 (declining to apply article 17 to a Muslim theologian’s texts as the Court was ‘unable to . . . conclude that the book in question incited violence, religious hatred or intolerance’); ECtHR, *Zemmour v. France* (App. no. 63539/19), 20 December 2022, §§28, 61, 64–67 (declining to apply article 17 to comments by a guest on a prime time television program describing Muslims in France as ‘invaders’ and ‘colonisers’ trying to ‘Islamize’ France despite the fact that they were ‘shocking,’ ‘controversial,’ ‘discriminatory’ and ‘likely to stir up a rift between the French and Muslim community,’ but approving his conviction and 3,000 Euro fine under article 10 read ‘in the light of Article 17’).

²⁸¹ IACtHR, *Kimel v. Argentina* (Series C, no. 177), 2 May 2008, §71.

²⁸² *Ibid.*

²⁸³ IACtHR, *Álvarez Ramos v. Venezuela* (Series C, no. 380), 30 August 2019, §107.

²⁸⁴ With the exception its decision regarding Rwanda’s denial and minimization of genocide laws: ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017 and its decision regarding Benin’s Digital Code which criminalizes ‘racially motivated and xenophobic insults using a computer system’ and ‘incitement to hate and violence’: ACtHPR, *Ajavon v. Republic of Benin* (App. no. 062/2019), 4 December 2020.

²⁸⁵ ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017, §133. See also ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §129; ACmHPR, *Good v. Botswana* (Comm. no. 313/05), 26 May 2010, §188.

²⁸⁶ ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §§134–135; ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017, §140; ECOWAS CCJ, *Federation of African Journalists v. The Gambia* (Suit no. ECW/CCJ/APP/36/15), 13 February 2018, 32–33; ECOWAS CCJ, *The Incorporated Trustees of Laws and Rights Awareness Initiatives v. Nigeria* (Suit no. ECW/CCJ/APP/53/2018), 10 July 2020, §§139–140. Similarly, the ECOWAS Community Court of Justice, which has jurisdiction over a number of western African states, has considered article 19(3) as providing the list of legitimate aims for restricting speech, either on its own or in addition to those recognized under article 27(2) of the African Charter: ECOWAS CCJ, *Federation of African Journalists v. The Gambia* (Suit no. ECW/CCJ/APP/36/15), 13 February 2018, 32–33; ECOWAS CCJ, *The Incorporated Trustees of Laws and Rights Awareness Initiatives v. Nigeria* (Suit no. ECW/CCJ/APP/53/2018), 10 July 2020, §§139–140.

compounded by the fact that the African Commission has cited article 27(2) as the ‘only legitimate reasons’ for limiting rights,²⁸⁷ and that the Commission’s 2019 Declaration of Principles on Freedom of Expression instead adopts the grounds enumerated in article 19(3), but excludes ‘morals.’²⁸⁸ The Commission has also said that ‘public interest’ can be recognized as a justifiable ground to limit expression under article 27(2), even though this is not listed in article 19(3).²⁸⁹

In one of the few cases in which the African Court has considered hate speech laws the Court cites a number of legitimate aims.²⁹⁰ The Court assessed whether Benin’s provisions which punish ‘the offences of racially motivated and xenophobic insults using a computer system’ and ‘incitement to hatred and violence on such grounds as race, colour, national or ethnic origin, or religion’ violated article 9(2) of the African Charter.²⁹¹ When considering whether this pursued a legitimate aim, the Court cited both article 27(2) and the fact that it had also ‘previously concluded that national security, public order and public morals are legitimate restrictions.’²⁹² It then held that the law ‘pursues a legitimate aim since it seeks to combat any form of incitement to hatred or discrimination.’²⁹³

Case Study: International Standards on Blasphemy Laws

The Human Rights Committee has found that certain aims are *not* legitimate bases for restricting speech, including laws that prohibit speech that is rude about religion or denies historic atrocities.²⁹⁴ In the context of blasphemy, for instance, the Committee has made clear that ‘[p]rohibitions of displays of lack of respect for a religion or other belief system, including blasphemy laws, are incompatible with the Covenant, except in the specific circumstances envisaged in article 20, paragraph 2.’²⁹⁵ The Committee has consequently recommended that states repeal their blasphemy laws.²⁹⁶ And in the case of *Rabbae v. The Netherlands*, it was noted that

²⁸⁷ ACmHPR, *Constitutional Rights Project v. Nigeria* (Comm. no. 140/1994), 15 November 1999, §41; ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §§134–135 (stating that ‘the only legitimate reasons’ for limiting rights are stipulated in article 27(2) while also ‘noti[ng] that the legitimate purpose of a restriction is stated in Article 19 (3) (a) and (b) of the Covenant’).

²⁸⁸ ACmHPR, Declaration of Principles of Freedom of Expression and Access to Information in Africa (2019), Principle 9.

²⁸⁹ ACmHPR, *Good v. Botswana* (Comm. no. 313/05), 26 May 2010, §189.

²⁹⁰ ACtHPR, *Ajavon v. Republic of Benin* (App. no. 062/2019), 4 December 2020. This case considered a number of other provisions of the Charter in addition to article 9.

²⁹¹ *Ibid.*, §121. The case was brought by a Beninese businessman and political refugee who had not been charged under Benin’s Digital Code, but could bring proceedings based on ‘a peculiarity of the African regional human rights system characterized by the objective nature of human rights litigation’ that meant he did not need to be a ‘victim’: ACtHPR, *Ajavon v. Republic of Benin* (App. no. 062/2019), 4 December 2020, §§58–59.

²⁹² ACtHPR, *Ajavon v. Republic of Benin* (App. no. 062/2019), 4 December 2020, §123 (citing ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017).

²⁹³ ACtHPR, *Ajavon v. Republic of Benin* (App. no. 062/2019), 4 December 2020, §123, §125.

²⁹⁴ See s. III.2.4.2. (Opinion).

²⁹⁵ HRC, General Comment No. 34 (2011), §48 (emphasis added).

²⁹⁶ See HRC, *Concluding Observations, Kuwait* (2016) UN Doc. CCPR/C/KWT/CO/3, §§39, 41 (‘The State party should eliminate all discriminatory legislation and practices that violate the right to freedom of thought, conscience and religion, including blasphemy laws that are incompatible with the Covenant.’); HRC, *Concluding Observation, Russian Federation* (2015) UN Doc. CCPR/C/RUS/CO/7, §19 (expressing concern about Russia’s

‘criticism of even the most deeply-held convictions of the adherents of a religion is protected by freedom of expression.’²⁹⁷ The Working Group on Arbitrary Detention has upheld this position as well.²⁹⁸ And the UN Special Rapporteur considers that blasphemy laws fail to pursue a legitimate aim under article 19(3) ‘given that Article 19 protects *individuals* and their rights to freedom of expression and opinion ... [It does not] protect *ideas or beliefs* from ridicule, abuse, criticism or other “attacks” seen as offensive.’²⁹⁹ In addition, under the Rabat Plan of Action, a codification of international standards supported by the UN, blasphemy laws are considered ‘counterproductive’ as ‘the right to freedom of religion or belief, as enshrined in relevant international legal standards, does not include the right to have a religion or a belief that is free from criticism or ridicule’.³⁰⁰

Similarly, the Inter-American Court has not tolerated censorship of speech simply on the basis that it offends religious sensibilities.³⁰¹ And some European bodies, including the Venice Commission—an expert body that forms part of the Council of Europe—the European Union and the Organization for Security and Co-operation in Europe, have followed this approach and advocated for the decriminalization of blasphemy laws.³⁰²

The European Court has, however, taken a different view and does not accept that blasphemy laws cannot constitute legitimate restrictions on speech. The Court has recognized a ‘wider margin of appreciation ... when regulating freedom of expression in relation to matters liable to offend intimate personal convictions within the sphere of ... religion’, and has held that, because what ‘is likely to cause substantial offence to persons of a particular religious persuasion will vary significantly from

‘blasphemy law’ and recommending Russia ‘repeal or revise’ its laws ‘with a view to bringing them into conformity with its obligations under the Covenant’). See also HRC, *Concluding Observations, Ireland* (2014) UN Doc. CCPR/C/IRL/CO/4, §22; HRC, *Concluding Observations, Indonesia* (2013) UN Doc. CCPR/C/IDN/CO/1, §25.

²⁹⁷ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, §10.5.

²⁹⁸ See, e.g., WGAD, *Mejri v. Tunisia* (Opinion no. 29/2013), 30 August 2013, §18; WGAD, *Amer v. Egypt* (Opinion no. 35/2008), 20 November 2008, §38. See also WGAD, *Ahmad v. Pakistan* (Opinion no. 7/2023), 3 May 2023, §74 (expressing ‘serious concern about blasphemy laws in Pakistan’). See s. II.2.2.3. (Pakistan).

²⁹⁹ UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §21 (emphasis added).

³⁰⁰ See Rabat Plan of Action, §19. For over a decade beginning in the late 1990s, the Organization of Islamic Cooperation (OIC) pushed for the adoption of a number of contentious resolutions in the UN Human Rights Council and at the UN General Assembly that recognized the concept of ‘defamation of religions’ and sought to either prohibit or criminalize speech disrespectful of religion, while Western nations consistently rejected the concept and the resolutions that sought to entrench it. In 2011, a consensus was reached by way of the Human Rights Council’s Resolution 16/18, which moved away from a legal prohibition regarding the defamation of religions and instead ‘presents an action-oriented approach to combat religious intolerance through practical steps that states should take, such as enforcing anti-discrimination laws and speaking out against intolerance, while also protecting freedoms of speech and religion’: see E. Aswad, R. Hussain, & A.M. Suleman, ‘Why the United States Cannot Agree to Disagree on Blasphemy Laws’ (2014) 32 *Boston University International Law Journal* 123, 144–145.

³⁰¹ See IACtHR, *The Last Temptation of Christ* (*Olmedo-Bustos et al.*) v. *Chile* (Series C, no. 73), 5 February 2001, §§71–73, 76–80 (holding that Chile’s prohibition of the exhibition of the film ‘The Last Temptation of Christ’ on the basis that it ‘deformed and diminished’ the image of Christ constituted prior censorship in violation of article 13, and did not violate article 12 (freedom of religion)).

³⁰² Venice Commission, *Report on the Relationship between Freedom of Expression and Freedom of Religion: The Issue of Regulation and Prosecution of Blasphemy, Religious Insult and Incitement to Religious Hatred* (2008) CDL-AD(2008)026, §92; EU Guidelines on the promotion and protection of freedom of religion or belief; OSCE, ‘OSCE Representative on Freedom of the Media welcomes Irish referendum on blasphemy’ (2 October 2017).

time to time and place to place ... State authorities are in principle in a better position than an international judge to give an opinion' on the need for such a restriction.³⁰³ According to the Court, although a 'religious group must tolerate the denial by others of their religious beliefs and even the propagation by others of doctrines hostile to their faith', this does not extend to statements that 'incite to hatred or religious intolerance'.³⁰⁴

The European Court has indeed gone as far as approving a criminal conviction for 'disparaging religious doctrines' in Austria after a woman referred to the 56-year-old Prophet Muhammad's marriage to a child as 'paedophilia'.³⁰⁵ The Court upheld the conviction on the basis that 'expressions that seek to spread, incite or justify hatred based on intolerance, including religious intolerance, do not enjoy the protection afforded by Article 10 of the Convention', and that the 'paedophile' comment was a 'value judgment without sufficient factual basis'³⁰⁶ and 'an abusive attack ... capable of stirring up prejudice and putting at risk religious peace'³⁰⁷ even though it was accepted that the Prophet married the girl when she was six years of age.

2.3. Necessity

The third arm of the tripartite test under article 19(3) of the ICCPR and regional free speech treaties governing when restrictions on speech are permissible is whether a penalty for speech is 'necessary'.³⁰⁸ The European Court has acknowledged that 'in the majority of cases, this is the question which determines the Court's conclusion in a given case', and the same applies to other international and regional bodies.³⁰⁹ The potential harm arising from hate speech, the causal link between speech and harm and the intent of the speaker are all relevant to an assessment of necessity. The requirement of 'necessity' also encompasses proportionality, as any restrictive measure 'must be the least intrusive instrument amongst those which might achieve their protective function' and must be 'proportionate to the interest to be protected'.³¹⁰

³⁰³ ECtHR, *Murphy v. Ireland* (App. no. 44179/98), 10 July 2003, §67. See also ECtHR, *Wingrove v. United Kingdom* (App. no. 17419/90), 25 November 1996, §53.

³⁰⁴ ECtHR, *Tagiyev v. Azerbaijan* (App. no. 13274/08), 5 December 2019, §44; ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §52. ECtHR, *Rabczewska v. Poland* (App. no. 8257/13), 15 September 2022, §57. See also ECtHR, *Yefimov and Youth Human Rights Group v. Russia* (App. nos. 12385/15 & 51619/15), 7 December 2021, §45 (finding no violation of article 10 when a speaker's 'criticism focused on the religious organization rather than on individual believers' and 'did not call for anyone's exclusion or discrimination, let alone incite to acts of violence or intimidation').

³⁰⁵ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §§43, 54. E.S. was charged under article 283 of the Austrian Criminal Code for inciting hatred. See s. II.2.1.2. (Austria).

³⁰⁶ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §§43–57. One expert has argued that the impact of this case will be that, '[g]iven the uncertainties that often surround historico-religious events, it may be impossible—or at least extremely challenging—to provide the factual basis the European Court now seems to require': S. Smet, 'Introductory Note to E.S. v. Austria' (2019) 58 *International Legal Materials* 628. The penalty of 480 Euros was also considered 'moderate' and therefore not disproportionate ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §56. See s. II.2.1.2. (Austria).

³⁰⁷ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §57.

³⁰⁸ See s. III. (International Legal Standards).

³⁰⁹ ECtHR, Guide on Article 10 of the European Convention on Human Rights—Freedom of Expression, 31 August 2022, §61.

³¹⁰ HRC, General Comment No. 34 (2011), §34.

2.3.1. Harm

In the leading case before a UN body related to hate speech, two members of the Human Rights Committee concluded that criminal penalties for hate speech should be limited to ‘speech that incites the commission of *criminal offences or acts of violence*’.³¹¹ And the Committee has found that ‘[w]hen a State ... invokes a legitimate ground for restrict[ing] freedom of expression, it must demonstrate in specific and individualized fashion the precise nature of the threat’ caused by speech and establish a ‘direct and immediate connection between the expression and the threat’.³¹² In addition, UN-approved international instruments such as the Rabat Plan of Action require ‘a *reasonable probability* that the speech would succeed in *inciting actual action* against the target group, recognizing that such causation should be rather direct’.³¹³ The ‘likelihood, including imminence’ of harm, is also one of the six threshold considerations identified as relevant to determining whether it is permissible to criminalize hate speech.³¹⁴

The European Court has taken a different approach. The potential harm that hate speech may cause is not the determinative factor in the European Court’s assessment of the necessity of restricting such speech, but one of a number of factors that are ‘taken into account’.³¹⁵ The Court has adopted a broad conception of what harm hate speech laws may lawfully seek to prevent. This includes civil or criminal penalties for statements that ‘could’, when read in context, ‘be seen as a call’ for ‘hatred or intolerance’.³¹⁶ In addition, in some of the hate speech cases that the European Court considers inadmissible under article 17, no assessment of harm is undertaken.³¹⁷ For instance, where the supporter of a far right political group possessed posters containing hateful messages such as ‘Romania needs children, not homosexuals’, his conviction was considered sound on the basis the posters ‘were intended to instigate hatred’ and ‘were

³¹¹ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §7 (emphasis added). This contrasts with an earlier decision by the Committee: HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996 which has since essentially been overruled. See also HRC, *Ross v. Canada* (Comm. no. 736/1997), 18 October 2000, §11.6 (regarding non-criminal penalties); HRC, General Comment No. 34 (2011), §49; M. O’Flaherty, ‘Freedom of Expression: Article 19 of the International Covenant on Civil and Political Rights and the Human Rights Committee’s General Comment No 34’ (2012) 12 *Human Rights Law Review* 627, 653 (describing General Comment No. 34 as ‘overruling *Faurisson*’); HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996, Individual Opinion of Committee members Elizabeth Evatt and David Kretzmer, co-signed by Eckart Klein, §9 (observing that the law used to convict the speaker did ‘not link liability to ... the tendency of the publication to incite to anti-semitism’).

³¹² HRC, General Comment No. 34 (2011), §35, citing HRC, *Shin v. Republic of Korea* (Comm. no. 926/2000), 16 March 2004. See also HRC, *Ross v. Canada* (Comm. no. 736/1997), 18 October 2000, §11.5.

³¹³ Rabat Plan of Action, §29(f) (emphasis added).

³¹⁴ *Ibid.* See also UN Special Rapporteur F. La Rue, *Promotion and Protection of the Right to Freedom of Opinion and Expression* (2013) UN Doc. A/68/362, §§52, 53.

³¹⁵ These factors include ‘the context in which the impugned statements were made, their nature and wording, their potential to lead to harmful consequences’, whether ‘the statements were made against a tense political or social background’ and whether they ‘could be seen as a direct or indirect call for violence or as a justification of violence, hatred or intolerance’. ECtHR, *Atamanchuk v. Russia* (App. no. 4493/11), 11 February 2020, §50. See also ECtHR, *Ibragimov v. Russia* (App. nos. 1413/08 & 28621/11), 28 August 2018, §99; ECtHR, *Alekhina v. Russia* (App. no. 38004/12), 17 July 2018, §221.

³¹⁶ ECtHR, *Ibragimov v. Russia* (App. nos. 1413/08 & 28621/11), 28 August 2018, §98; citing ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §206.

³¹⁷ See, e.g., ECtHR, *Ivanov v. Russia* (App. no. 35222/04), 20 February 2007. See also ECtHR, *Norwood v. United Kingdom* (App. no. 23131/03), 16 November 2004. See ss III.2.4.2. (Opinion) and III.2. (Discretionary Restrictions on Hate Speech).

likely to seriously disturb public order’ even though the posters were never publicly displayed and these harms were merely hypothetical.³¹⁸

The European Court considers that it may be necessary ‘to sanction or even prevent all forms of expression which spread, incite, promote or justify hatred based on intolerance (including religious intolerance),’ provided that restrictions are proportionate.³¹⁹ Speech does not need to ‘involve an explicit call for an act of violence, or other criminal acts’ for a penalty to be considered ‘necessary,’ as states may choose to validly restrict speech ‘insulting, holding up to ridicule or slandering specific groups of the population.’³²⁰ The Court has applied this reasoning to speech targeting a range of different groups, including racist speech, religious intolerance, ‘extreme nationalism’ and discrimination on the basis of sexual orientation.³²¹

The European Court has found that hate speech could be penalized even when the words used could be ‘interpreted in several ways’ and were contradictory.³²² For instance, France was held to have legitimately convicted and fined a cartoonist for publishing satirical drawings of the September 11, 2001 attacks with the caption ‘We have all dreamt of it . . . Hamas did it.’³²³ Since it was printed two days after the attacks, ‘when the whole world was in a state of shock from the news’ and ‘in a politically sensitive [Basque] region’ of France, the Court held that the cartoon ‘elicited reactions which *could* have stirred up violence and which demonstrated its plausible impact on public order in the region.’³²⁴ Commentators have, however, pointed to the small circulation of the local newspaper and the differences between terrorism in the United States and the Basque region and concluded that ‘it is hard to maintain that the drawing in question created a ‘credible’ danger of violence, and that the case demonstrates that the ‘capable of’ causing harm test ‘is a vague one that allows judges considerable scope to prohibit speech merely because they consider it to be highly offensive.’³²⁵

Despite states’ wide ‘margin of appreciation,’³²⁶ and the reliance on article 17, the European Court has sometimes found that a penalty for hate speech was not necessary

³¹⁸ ECtHR, *Molnar v. Romania* (App. no. 16637/06), 23 October 2012, §23 (some speech was also anti-Roma, and there was evidence before domestic courts that the applicant had previously displayed far-right posters publicly). See also ECtHR, *Lilliendahl v. Iceland* (App. no. 29297/18) 12 May 2020, §36.

³¹⁹ ECtHR, *Ibragimov v. Russia* (App. nos. 1413/08 and 28621/11), 28 August 2018, §94.

³²⁰ *Ibid.* See also ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012, §55.

³²¹ See ECtHR, *Ibragimov v. Russia* (App. nos. 1413/08 & 28621/11), 28 August 2018, §94 (observing that the Court ‘has been particularly sensitive towards sweeping statements attacking or casting in a negative light entire ethnic, religious or other groups’); ECtHR, *Balsytė-Lideikienė v. Lithuania* (App. no. 72596/01), 4 November 2008, §80 (finding no violation of article 10 where statements ‘promoted national hatred, xenophobia and territorial claims’); ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012, §55 (holding that ‘discrimination based on sexual orientation is as serious’ as discrimination based on ‘race, origin or colour’).

³²² ECtHR (GC), *Zana v. Turkey* (App. no. 18954/91), 25 November 1997, §§58–62. See also ECtHR, *Atamanchuk v. Russia* (App. no. 4493/11), 11 February 2020, §64.

³²³ ECtHR, *Leroy v. France* (App. no. 36109/03), 2 October 2008, §§6, 48.

³²⁴ *Ibid.*, §45 (emphasis added).

³²⁵ S. Sottiaux, ‘Leroy v. France: Apology of Terrorism and the Malaise of the European Court of Human Rights’ Free Speech Jurisprudence’ (2009) 3 *European Human Rights Law Review* 415, 424. See ch. 6 (Terrorism Laws), s. III.3.2. (Causal link between speech and harm). The Court will often consider the manner or medium in which a hateful statement is made to determine whether its capacity to lead to harmful consequences. See ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §207; ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012, §56.

³²⁶ See s. III.2. (Discretionary Restrictions on Hate Speech).

or proportionate because domestic courts failed to provide ‘relevant and sufficient’ reasons to justify the necessity of a penalty for speech. For example, when a writer and editor were convicted of incitement to religious hatred and hostility for publishing an article that criticized Islam by describing the Prophet as ‘a frightful creature’, the European Court found that the Azerbaijan courts had not explained ‘why the remarks . . . constituted incitement to religious hatred and hostility’ and that the speaker’s rights had been violated on this basis.³²⁷

The approach of the Inter-American system is closer to that of the Human Rights Committee than the European Court. Its Rapporteur has held that states must demonstrate ‘both the existence of an impending threat that [the speech] could cause real harm and that it is crucial to impose the restriction in order to prevent the harm.’³²⁸ In order to justify restrictions on freedom of expression to protect the rights of others, those rights must ‘be clearly harmed or threatened.’³²⁹ A speaker must have had both intent to ‘promot[e] lawless violence or similar action’ and ‘the capacity to achieve this objective and create an *actual* risk of harm.’³³⁰

Similarly, guidance from the African Commission provides that penalties for hate speech must be based on a ‘real likelihood and imminence of harm’ and ‘a close causal link between the risk of harm and the expression.’³³¹ And in a short decision considering Benin’s hate speech laws, the African Court held that in ‘view of the harmful consequences such rhetoric can engender’ a criminal penalty in such laws was not necessarily ‘disproportionate given their deterrent function.’³³²

2.3.2. *Intent*

Unlike the clear intent requirements in the context of mandatory penalties for speech,³³³ international and regional bodies have not been explicit when addressing intention in relation to discretionary restrictions of hate speech.

In the context of defamation laws, the Human Rights Committee has stated that ‘consideration should be given to avoiding penalizing or otherwise rendering unlawful untrue statements that have been published in error but *without malice*.’³³⁴ This endorses a very high standard of intent, codified at the domestic level in US federal law in the

³²⁷ ECtHR, *Tagiyev v. Azerbaijan* (App. no. 13274/08), 5 December 2019, §47. The Court also found that the writer and editor’s terms of imprisonment, being three and four years respectively, were disproportionate sanctions. See also ECtHR, *Ibragimov v. Russia* (App. nos. 1413/08 & 28621/11), 28 August 2018, §106; ECtHR, *Stomakhin v. Russia* (App. no. 52273/07), 9 May 2018, §129–132.

³²⁸ IACmHR, *Annual Report of the Inter-American Commission on Human Rights: Report of the Special Rapporteur for Freedom of Expression* (2009) OEA/Ser.L/V/II., Doc. 51, 397.

³²⁹ *Ibid.*, 250.

³³⁰ IACmHR, *Violence against Lesbian, Gay, Bisexual, Trans and Intersex Persons in the Americas* (2015) OAS/Ser.L/V/II. Doc.36/15 Rev.2, §235 (in the context of mandatory restrictions on speech).

³³¹ See ACmHPR, *Declaration of Principles of Freedom of Expression and Access to Information in Africa* (2019), Principle 22(5) (related to restrictions on grounds of public order or national security), Principle 23(2)(f).

³³² ACtHPR, *Ajalon v. Republic of Benin* (App. no. 062/2019), 4 December 2020, §127. See s. III.2.2. (Legitimacy).

³³³ See s. III.1. (Mandatory Restrictions on Hate Speech).

³³⁴ HRC, General Comment No. 34 (2011) (emphasis added). Cf., in the context of denial laws; HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996, Individual Concurring Opinion of Committee members Elizabeth Evatt and David Kretzmer, co-signed by Eckart Klein, §9 stating that the restrictions in that case ‘do not meet the proportionality test. They do not link liability to the intent of the author, nor to the tendency of the publication to incite anti-Semitism.’

seminal Supreme Court case of *New York Times v. Sullivan*.³³⁵ In addition, in the context of hate speech laws, the UN Special Rapporteur has taken the position that '[n]o one should be penalized for the dissemination of hate speech unless it has been shown that they did so with the *intention of inciting* discrimination, hostility or violence.'³³⁶

Although the European Court considers intent a relevant factor in determining the necessity of restricting hate speech,³³⁷ in a number of cases the Court has, even in the criminal context, allowed convictions without a clear showing of intent or simply implied intent on the basis of the content of the speech. For example, a cartoonist convicted for his cartoon of the September 11, 2001 attacks argued that he did not intend to encourage terrorism but to express his anti-American sentiment.³³⁸ The Court, however, found no violation of article 10, noting that the drawing could 'in itself demonstrate the intention of the author', but that even when taken together with the accompany caption, 'the work does not criticize American imperialism, but supports and glorifies its destruction by violence'. The Court further adopted the domestic court's finding that 'the applicant's intentions are unrelated to the prosecution', given that his intentions had only been expressed retroactively and 'were not of such a nature, in view of the context, to erase [his] positive assessment of the consequences of a criminal act'.³³⁹ In other decisions, speech restrictions have been justified even though the Court accepted that the speaker *did not* have intent to create a specific harm. For example, when homophobic leaflets were distributed at a school, the Court found that the conviction and fine against those who distributed the leaflets did not violate article 10, even though they 'had not intended to express contempt for homosexuals as a group' but to 'start a debate about the lack of objectivity in the education dispensed in Swedish schools'.³⁴⁰

The Inter-American Commission on Human Rights, adopting a position similar to US domestic law,³⁴¹ has held that hate speech sanctions can only be justified when a speaker was 'not simply issuing an opinion' but also had 'intent' to 'promot[e] lawless violence or similar action'.³⁴² And in the African context, the Declaration of Principles on Freedom of Expression in Africa provides that when determining if speech falls into

³³⁵ US Supreme Court, *New York Times v. Sullivan* 376 U.S. 254, 9 March 1964.

³³⁶ UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §50 (emphasis added).

³³⁷ See, e.g., ECtHR (GC), *Jersild v. Denmark* (App. no. 15890/89), 23 September 1994, §33 ('the object of the programme was to address aspects of the problem, by identifying certain racist individuals and by portraying their mentality and social background. There is no reason to doubt that the ensuing interviews fulfilled that aim. Taken as a whole, the feature could not objectively have appeared to have as its purpose the propagation of racist views and ideas').

³³⁸ ECtHR, *Leroy v. France* (App. no. 36109/03), 2 October 2008, §42. See s. III.2.3.1. (Harm).

³³⁹ *Ibid.*, §43. See also ECtHR (GC), *Sürek v. Turkey (No. 1)* (App. no. 26682/95), 8 July 1999, §62.

³⁴⁰ ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012, §10. See also ECtHR, *Nix v. Germany* (App. no. 35285/16), 13 March 2018, §51 (finding a criminal conviction and fine inadmissible even where '[t]he Court accepts that the applicant, by displaying the picture of Himmler in SS uniform with a swastika armband in his blog post, did not intend to spread totalitarian propaganda, to incite violence, or to utter hate speech, and that his expression had not resulted in intimidation').

³⁴¹ See s. II.2.3.2. (United States). See ch. 2 (Insulting Speech), s. II.1.2. (Intent) (regarding the adoption of the 'actual malice' standard in the Inter-American context).

³⁴² IACmHR, *Violence against Lesbian, Gay, Bisexual, Trans and Intersex Persons in the Americas* (2015) OAS/Ser.L/V/II. Doc.36/15 Rev.2, §235.

the exceptional category of warranting criminal sanctions, states should take into account the ‘existence of a clear intent to incite.’³⁴³

2.4. Exclusions, exceptions and defences

2.4.1. Truth

The Human Rights Committee has made clear that, at least in the context of criminal laws that penalize defamatory speech, courts should recognize a ‘defence of truth’ and that such criminal laws ‘should not be applied with regard to those forms of expression that are not, of their nature, subject to verification.’³⁴⁴ In addition, the UN Special Rapporteur has recommended a number of limits on hate speech laws, including that ‘no one should be penalized for statements that are true.’³⁴⁵ Similarly, the Declaration of Principles on Freedom of Expression in Africa 2019 provides that, at least in relation to defamation laws, ‘[n]o one shall be found liable for true statements’.

The European Court, on the other hand, has not recognized a defence of truth in hate speech cases.³⁴⁶ In some instances, the Court has simply asserted the falsity of a statement and considered this relevant to finding its penalization valid.³⁴⁷ For instance, in a case in which a woman argued that her statement that the Prophet Muhammad was a ‘paedophile’ was true since the Prophet married a six-year-old girl,³⁴⁸ the European Court disagreed and instead accepted the Austrian court’s assessment that these statements were ‘value judgments without sufficient factual basis.’³⁴⁹ And it found that even if the speech was factual, the speaker had ‘failed to adduce any evidence’ of its truth and ‘must have been aware that her statements were partly based on untrue facts.’³⁵⁰

2.4.2. Opinion

The Human Rights Committee recognizes a defence of opinion in the context of defamation law.³⁵¹ And in relation to denial laws, the Human Rights Committee has concluded that laws ‘that penalize *the expression of opinions* about historical facts are incompatible’ with ‘respect for freedom of ... expression’, and that the ICCPR ‘does not permit general prohibition of expressions of an erroneous *opinion* or an incorrect interpretation of past events.’³⁵² This position represents a significant reversal of the

³⁴³ ACmHPR, Declaration of Principles of Freedom of Expression and Access to Information in Africa (2019), Principle 23(2)(c).

³⁴⁴ HRC, General Comment No. 34 (2011), §47.

³⁴⁵ UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §50 (citing the UN Special Rapporteur on Freedom of Opinion and Expression, the OSCE Representative on Freedom of the Media and the OAS Special Rapporteur on Freedom of Expression, Joint Statement on Racism and the Media (2001)).

³⁴⁶ The position is different when it comes to defamatory speech. See ch. 2 (Insulting Speech), s. III.4.1. (Truth).

³⁴⁷ ECtHR, *Pastörs v. Germany* (App. no. 55225/14), 3 October 2019, §48. Cf. ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §243.

³⁴⁸ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §34.

³⁴⁹ *Ibid.*, §54. See s. III.2.2. (Legitimacy) (Case Study: International Standards on Blasphemy Laws).

³⁵⁰ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §§53–55. See also S. Smet, ‘Case note: Free Speech versus Religious Feelings, the Sequel: Defamation of the Prophet Muhammad in *E.S. v Austria*’ (2019) 15 *European Constitutional Law Review* 158, 166–167. See ss II.2.1.2. (Austria) and VI. (Recommendations).

³⁵¹ See ch. 2 (Insulting Speech), s. II.4.5. (Opinion).

³⁵² HRC, General Comment No. 34 (2011), §49.

Committee's earlier approach to this issue,³⁵³ and the Committee has since criticized so-called 'denial laws' in a number of countries.³⁵⁴ The Human Rights Committee has, however, suggested that some restrictions to freedom of expression regarding historical facts could be permitted if they constitute speech that *must be* penalized under article 20 of the ICCPR.³⁵⁵

The CERD Committee has taken a similar position, recommending that 'the expression of opinions about historical facts' should not be prohibited but acknowledging that 'public denials or attempts to justify crimes of genocide and crimes against humanity ... should be declared as offences punishable by law' if 'they clearly constitute incitement to racial violence or hatred'.³⁵⁶ And the Inter-American Commission has suggested more broadly that states seeking to prohibit hate speech inciting violence 'must have as a prerequisite strong, objective evidence that the person was not simply expressing an opinion, but also had the clear intention to commit an unlawful act'.³⁵⁷

The European Court has taken a different approach and consistently held that stating certain opinions—such as a belief that the Holocaust did not take place—*cannot be* protected speech, and the Court frequently strikes out claims relating to such speech on the basis that they run counter to the 'fundamental values' under article 17 of the Convention.³⁵⁸

For instance, when a French philosopher published a book that disputed the commission of crimes against humanity during the Holocaust and was sentenced to a three-year suspended prison sentence and a fine, the Court held the case inadmissible because his work was not 'historical research akin to a quest for truth,' with its 'real purpose being to rehabilitate the National-Socialist regime,' thereby running counter to

³⁵³ M. O'Flaherty (n 311) 627, 653 (referring to HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996). The professor was convicted and fined under France's Gayssot Act. The Committee held that the law sought to ensure the Jewish community could 'live free from fear of an atmosphere of anti-semitism' and that the restriction to the professor's speech was necessary since 'the denial of the existence of the Holocaust' was a 'principle vehicle for anti-semitism': HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996, §§9.6–9.7. See also Individual Opinion of Committee members Elizabeth Evatt and David Kretzmer, co-signed by Eckart Klein, §10 (stating that while there is 'every reason to maintain protection of bona fide historical research against restriction, even when it challenges accepted historical truths and by so doing offends people, anti-semitic allegations of the sort made by the author ... do not have the same claim to protection'). Cf. HRC, *Zündel v. Canada* (Comm. no. 953/2000), 27 July 2003, §8.5.

³⁵⁴ See s. III.1.1. (ICCPR Article 20). See, e.g., HRC, *Concluding Observations, Lithuania* (2018) UN Doc. CCPR/C/LTU/CO/4, §§27–28 (expressing concern 'at reports of recently proposed legislative amendments that would ban the sale of material that "distorts historical facts" about the nation'); HRC, *Concluding Observations, Russia* (2015) UN Doc. CCPR/C/RUS/CO/7, §19 (expressing concern at a law criminalizing 'distortion of the Soviet Union's role in the Second World War' and suggesting Russia 'repeal or revise' this and other laws to bring them into conformity with the ICCPR). See also HRC, *Concluding Observations, Poland* (2016) UN Doc. CCPR/C/POL/CO/7, §38.

³⁵⁵ HRC, General Comment No. 34 (2011), §49.

³⁵⁶ CERD Committee, General Recommendation No. 35 (2013), §14 (citing to HRC, General Comment No. 34 (2011), §49).

³⁵⁷ IACmHR, *Annual Report of the Inter-American Commission on Human Rights: Report of the Special Rapporteur for Freedom of Expression* (2009) OEA/Ser.L/V/II., Doc. 51, §544. See also IACtHR, *Kimel v. Argentina* (Series C, no. 177), 2 May 2008, §93 (stating that 'an opinion cannot be subjected to sanctions' and that the 'truthfulness or falseness' of a statement 'may only be established in respect of facts'); IACtHR *Palacio Urrutia et al. v. Ecuador* (Series C, no. 446), 24 November 2021, §115.

³⁵⁸ ECtHR, *Pastörs v. Germany* (App. no. 55225/14), 3 October 2019, §§36–37. See also ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §115.

the Convention's values of 'justice and peace'.³⁵⁹ The Court also held that—even though any restriction to statements made 'by elected representatives in Parliament' requires the 'closest scrutiny'—a parliamentarian who denied the Holocaust in a speech in Germany legitimately received an eight-month suspended sentence.³⁶⁰ And the Court considered that a comedian who brought a Holocaust-denying academic on stage to receive an 'award' from an actor dressed as a Jewish deportee in a concentration camp could be convicted of racial insults without any review by the Court.³⁶¹

The Court's approach to Holocaust denial may be something of an aberration that results from the fact that its membership is made up of states that were implicated in the events of the Second World War.³⁶² Although this historical significance is understandable, such a viewpoint-based approach to regulating speech, particularly through the criminal law, is an unwelcome legal precedent. It has also led to unprincipled line-drawing, for instance in a case in which the Court reviewed a conviction in Switzerland for the statement that 'there was no genocide of the Armenians in 1915'. The Court found that the Swiss law used to convict the speaker stood 'at one end of the comparative spectrum' by criminalizing the denial of any genocide, 'without the requirement that it be carried out in a manner likely to incite to violence or hatred' and that its application to the speaker violated his rights.³⁶³ It therefore made the question of whether genocide denial is protected dependent on which genocide was involved.

Although the African Court has generally interpreted the right to free expression in line with the approach of the Human Rights Committee, it considers—like the European Court but unlike the Committee—that opinions about historical facts can be appropriately penalized.³⁶⁴ The Court reached this conclusion in a case in which a Rwandan politician was sentenced to 15 years' imprisonment for a number of offences, including 'minimization' of genocide, after she expressed concern that a memorial 'only refers to the people who died during the genocide against the Tutsis' and not the 'untold story with regard to the crimes against humanity committed against the Hutus'.³⁶⁵ The Court observed that although her statement did not 'deny or belittle' the genocide,³⁶⁶ 'it [was] entirely legitimate for the State to have introduced laws on the "minimisation", "propagation" and "negation" of genocide' and that statements which 'deny or minimize the magnitude or effects of the genocide or that unequivocally insinuate the same

³⁵⁹ ECtHR, *Garaudy v. France* (App. no. 65831/01), 24 June 2003. Five separate criminal proceedings were brought against the author as a result of the book, with a number of suspended sentences being handed down, to be served concurrently with the longest in time being three years.

³⁶⁰ ECtHR, *Pastörs v. Germany* (App. no. 55225/14), 3 October 2019, §§46–49.

³⁶¹ ECtHR, *M'Bala M'Bala v. France* (App. no. 25239/13), 20 October 2015 (inadmissible under Art. 17) (noting that although article 17 has 'always been applied to explicit and direct remarks not requiring any interpretation', in this case that 'the blatant display of a hateful and anti-Semitic position disguised as an artistic production [was] as dangerous as a fully-fledged and sharp attack').

³⁶² See s. III.2.2. (Legitimacy).

³⁶³ ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §§255–257. See s. III.2.2. (Legitimacy).

³⁶⁴ See s. III.2.2. (Legitimacy).

³⁶⁵ ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017, §§151, 158. The applicant was also convicted of other offences such as 'conspiracy to undermine established authority and violate constitutional principles by resorting to terrorism and armed force': §23. Some alleged offences were not based on speech.

³⁶⁶ *Ibid.*, §158.

fall outside the domain of the legitimate exercise of the right to freedom of expression and should be prohibited by law.³⁶⁷

2.4.3. Public interest

The Human Rights Committee has made clear that, in the context of defamation laws, a ‘public interest in the subject matter of a criticism’ should be recognized as a defence.³⁶⁸ The Inter-American and African Courts have similarly expressed a high threshold for restricting speech that deals with matters of public interest or relates to public officials.³⁶⁹

Similarly, the European Court has recognized that there is ‘little scope’ under article 10(2) ‘for restrictions to political speech or debates on questions of public interest’, and it ‘has been the Court’s consistent approach to require very strong reasons for justifying restrictions on political debate’, establishing a ‘particularly narrow margin of appreciation’ in such cases.³⁷⁰ At the other end of the spectrum, when it comes to speech that ‘incite[s] to violence’, the state authorities ‘enjoy a wider margin of appreciation when examining the need for an interference with freedom of expression’, and the same applies to speech ‘liable to offend intimate personal convictions within the sphere of morals or, especially, religion.’³⁷¹ As the Court has put it: ‘expression on matters of public interest is in principle entitled to strong protection, whereas expression that promotes or justifies violence, hatred, xenophobia or another form of intolerance cannot normally claim protection’ at all.³⁷² As a result, the Court has concluded that a state can impose criminal sanctions for speech ‘on an issue of legitimate public interest ... only in exceptional circumstances, notably where other fundamental rights have been seriously impaired’ but such circumstances include ‘for example’ both ‘incitement to violence’ and ‘hate speech.’³⁷³

Speech determined to be in the public interest in this context includes a documentary describing the attitudes of members of a racist youth group in Denmark and the band Pussy Riot performing their song ‘Punk Prayer—Virgin Mary, Drive Putin Away’ in a church to highlight ‘the political situation in Russia and the stance of [members of the Russian Orthodox Church] towards street protests in a number of Russian cities.’³⁷⁴ And when two protesters set fire to a photograph of the Spanish royal couple during a rally, the Court considered that this act was ‘part of a debate on issues of public interest,

³⁶⁷ *Ibid.*, §§147, 158. The Court found that the impugned speech did not amount to denying genocide, and that even ‘if this Court were to accept that there was a need to put restrictions on such statements’ a 15-year prison term ‘was not proportionate to the legitimate purposes which the conviction and sentence seek to achieve’: ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017, §162.

³⁶⁸ HRC, *Kozlov v. Belarus* (Comm. no. 1986/2010), 24 July 2014, §7.5–7.6.

³⁶⁹ See ch. 2 (Insulting speech), s. II.4.2. (Public interest). See, e.g., IACtHR, *Herrera-Ulloa v. Costa Rica* (Series C, no. 107), 2 July 2004, §101(2)(c); ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §155.

³⁷⁰ ECtHR, *Alekhina v. Russia* (App. no. 38004/12), 17 July 2018, §212; ECtHR (GC), *Bédat v. Switzerland* (App. no. 56925/08), 29 March 2016, §49. Cf. ECtHR (GC), *Sanchez v. France* (App. no. 45581/15), 15 May 2023, §176.

³⁷¹ ECtHR, *Murphy v. Ireland* (App. no. 44179/98), 10 July 2003, §67; ECtHR (GC), *Sürek v. Turkey* (App. nos. 23927/94 and 24277/94), 8 July 1999, §34.

³⁷² ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §230.

³⁷³ ECtHR, *Savva Terentyev v. Russia* (App. no. 10692/09), 28 August 2018, §83.

³⁷⁴ ECtHR (GC), *Jersild v. Denmark* (App. no. 15890/89), 23 September 1994, §33; ECtHR, *Alekhina v. Russia* (App. no. 38004/12), 17 July 2018, §212.

namely Catalan independence, the monarchical form of the State and criticism of the King as a symbol of the Spanish nation.³⁷⁵

The European Court has previously considered ‘the role of religion in society and its role in the development of society’ to be a matter of public interest, but overrides this principle when it considers that such speech ‘incite[s] to hatred or religious intolerance.’³⁷⁶ For instance, when a woman argued that her statement regarding the marriage of the Prophet Muhammad to a six-year-old had been ‘an objective criticism of religion’ and had ‘contributed to a public debate,’ the Court endorsed the domestic court’s view that these statements ‘had not been made in an objective manner aiming at contributing to a debate of public interest, but could only be understood as having been aimed at demonstrating that Muhammad was not a worthy subject of worship.’³⁷⁷ On the other hand, the Court held that the denial of the Armenian genocide by a politician was an issue of public interest, as ‘part of a long-standing controversy’ and a ‘heated debate, not only within Turkey but also in the international arena’ on the issue.³⁷⁸ In that same decision, the Court also noted that ‘statements on historical issues’ are ‘as a rule seen as touching upon matters of public interest,’ although the Court has never considered speech questioning the Holocaust to be in this category.³⁷⁹

2.4.4. ‘Responsible journalism’

The European Court has held that when it comes to assessing the necessity of speech restrictions, special protections may apply to the press.³⁸⁰ According to the Court, the doctrine of ‘reasonable journalism’ means that there should be ‘increased protection’ afforded to speech by the media and other ‘public watchdogs’ that have a duty to report on ‘matters of public interest’ as long as the reporting is objectively ‘reasonable.’³⁸¹ Since the Court has held that ‘public interest’ may relate to the ‘role of a religion in society and its role in the development of society’³⁸² and ‘statements on historical issues,’³⁸³ this principle may apply in the context of hate speech.

³⁷⁵ ECtHR, *Stern Taulats and Roura Capellera v. Spain* (Apps. no. 51168/15 & 51186/15), 13 March 2018, §36. The Court was not convinced that this could be regarded as ‘incitement to hatred or violence’ and found the protesters’ convictions and 2,700 Euro fines to be a violation of article 10: §§40–42.

³⁷⁶ ECtHR, *Tagiyev v. Azerbaijan* (App. no. 13274/08), 5 December 2019, §§44–45.

³⁷⁷ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §§34, 52. See ss II.2.1.2. (Austria) and III.2.2. (Legitimacy) (Case Study: International Standards on Blasphemy Laws).

³⁷⁸ ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §231. See ss III.2.2. (Legitimacy) and III.2.4.2. (Opinion).

³⁷⁹ ECtHR, *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §230. See s. III.2.4.2. (Opinion).

³⁸⁰ The IACtHR has quoted the approach of the ECtHR that journalists ‘must, when exercising their duties, abide by the principles of responsible journalism, namely to act in good faith, provide accurate and reliable information, objectively reflect the opinions of those involved in a public debate, and refrain from pure sensationalism.’ But the doctrine of ‘[r]esponsible journalism’ appears to operate more as a general principle rather than a defence or a way of assessing whether a violation of the right to free expression has occurred: IACtHR, *Mémoli v. Argentina* (Series C, no. 265), 22 August 2013, §122; see ch. 2 (Insulting Speech), s. II.1.4.3. (Reasonable publication).

³⁸¹ ECtHR, Guide on Article 10 of the European Convention on Human Rights—Freedom of Expression, 31 August 2022, §305. See also ECtHR (GC), *Fressoz v. France* (App. no. 29183/95), 21 January 1999, §45(ii); ECtHR (GC), *Pedersen v. Denmark* (App. no. 49017/99), 17 December 2004, §71. The responsible journalism doctrine applies even where journalistic reporting does not concern matters in the public interest: ECtHR (GC), *Satakunnan Markkinapörssi Oy v. Finland* (App. no. 931/13), 27 June 2017, §§174–183. Cf. Dissenting Opinion of Judges Sajó and Karakas, §2.

³⁸² ECtHR, *Tagiyev v. Azerbaijan* (App. no. 13274/08), 5 December 2019, §45.

³⁸³ ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §230. See s. III.2.4.2. (Opinion).

The Court has found that when freedom of the press is at stake, authorities have a limited margin of appreciation to decide that a ‘pressing social need’ for restrictions exists.³⁸⁴ But the Court has also made clear that journalists do not have an unfettered right to claim immunity from criminal liability if an offence was committed during the performance of journalistic activities.³⁸⁵ This protection is, according to the Court, therefore ‘subject to the condition that they comply with the duties and responsibilities connected with the function of journalist, and the consequent obligation of “responsible journalism”’.³⁸⁶ These duties and responsibilities ‘assume an even greater importance in situations of conflict and tension.’³⁸⁷ And this conditional protection has also been applied to other speakers, such as human rights activists, NGOs and whistleblowers who engage in public debate.³⁸⁸ The European Court has held, in a case involving a journalist, that the Court ‘will examine whether the journalist ... acted in good faith in accordance with the tenets of responsible journalism’.³⁸⁹ Whether or not a journalist acted ‘responsibly’ is assessed on the basis of the content of the information which is collected and disseminated by journalistic means as well as the process of publication.³⁹⁰

For example, in a case concerning a documentary airing the views of a racist group of young people in Denmark known as the ‘Greenjackets’, the Court observed that ‘methods of objective and balanced reporting may vary considerably’ and that ‘[i]t is not for [the] Court, nor for the national courts for that matter, to substitute their own views for those of the press as to what technique of reporting should be adopted by journalists’.³⁹¹ However, the Court had regard to ‘the manner in which the Greenjackets feature was prepared, its contents, the context in which it was broadcast and the purpose of the programme’, and found that taken as a whole, the feature sought ‘to expose, analyse and explain this particular group of youths’ rather than propagate racist views.³⁹² As a result, even though the remarks made by the Greenjackets themselves were not protected, those of journalists were.³⁹³

³⁸⁴ ECtHR, Guide on Article 10 of the European Convention on Human Rights—Freedom of Expression, 31 August 2022, §299. See ch. 1 (Introduction), s. II.3.1.5. (Importance of protecting the press).

³⁸⁵ ECtHR, Guide on Article 10 of the European Convention on Human Rights—Freedom of Expression, 31 August 2022, §320; ECtHR (GC), *Stoll v. Switzerland* (App. no. 69698/01), 10 December 2007, §102. See also ECtHR, *De Haes v. Belgium* (App. no. 19983/92), 24 February 1997, §37; ECtHR (GC), *Bladet Tromsø v. Norway* (App. no. 21980/93), 20 May 1999, §62.

³⁸⁶ ECtHR, Guide on Article 10 of the European Convention on Human Rights—Freedom of Expression, 31 August 2022, §380. See, e.g., ECtHR, *Kaçki v. Poland* (App. no. 10947/11), 4 July 2017, §49.

³⁸⁷ ECtHR, *Stomakhin v. Russia* (App. no. 52273/07), 9 May 2018, §102.

³⁸⁸ ECtHR (GC), *Magyar Helsinki Bizottság v. Hungary* (App. no. 18030/11), 8 November 2016, §159; ECtHR, *Gawlik v. Liechtenstein* (App. no. 23922/19), 16 February 2021, §77.

³⁸⁹ ECtHR, *Kaçki v. Poland* (App. no. 10947/11), 4 July 2017, §49 (emphasis added). The Court has also asked whether the journalist ‘act[ed] in good faith, respect[ed] the ethics of journalism and perform[ed] the due diligence expected in responsible journalism?’ ECtHR, *Magyar Jeti Zrt v. Hungary* (App. no. 11257/16), 4 December 2018, §77.

³⁹⁰ See ch. 1 (Introduction), s. II.3.1.5. (Importance of protecting the press).

³⁹¹ ECtHR (GC), *Jersild v. Denmark* (App. no. 15890/89), 23 September 1994, §31. See also ECtHR (GC), *Bladet Tromsø v. Norway* (App. no. 21980/93), 20 May 1999, §63.

³⁹² ECtHR (GC), *Jersild v. Denmark* (App. no. 15890/89), 23 September 1994, §§31, 33. This case also developed the Court’s jurisprudence providing that ‘the punishment of a journalist for assisting in the dissemination of statements made by another person in an interview would seriously hamper the contribution of the press to discussion of matters of public interest and should not be envisaged unless there are particularly strong reasons for doing so’: §35.

³⁹³ *Ibid.*, §§35–37. See ch. 1 (Introduction), s. II.3.1.5. (Importance of protecting the press).

2.5. Penalties

An assessment of whether hate-speech restrictions are ‘necessary’ to protect a legitimate aim under international human rights law includes an analysis of the proportionality of the penalty imposed.³⁹⁴ International courts and human rights bodies have made clear that criminal penalties are only justified in exceptional circumstances. However, the European Court defines such circumstances in broader terms than its counterparts and its jurisprudence is an outlier in finding that criminal convictions—including prison terms—for hate speech are compatible with the right to freedom of speech.³⁹⁵

In the small number of hate speech cases that have come before the Human Rights Committee, the Committee has recognized that even minor civil penalties can have a chilling effect on speech. And it has never approved of criminal penalties for speech except in one case involving a Korean election law (which did not involve a prison term) and a now-overturned decision on Holocaust-denial laws.³⁹⁶ In a concurring opinion in *Rabbae v. The Netherlands*, the leading hate speech case, two Committee members urged ‘great caution in the imposition of criminal penalties that punish speech’, proposing that criminal penalties should be limited to speech ‘that incites the commission of criminal offences or acts of violence’, particularly in recognition of the fact that hate speech laws ‘are often employed to suppress the very minorities they purportedly are designed to protect.’³⁹⁷ The UN has also recommended a range of non-censorial means to combat intolerance without resorting to criminal penalties against speech, such as inter-faith dialogue, education and other policies to combat intolerance.³⁹⁸

The European Court, for its part, considers that it may be ‘necessary in democratic societies to sanction or even prevent all forms of expression which spread, incite, promote or justify violence or hatred based on intolerance’ provided that penalties imposed ‘are proportionate.’³⁹⁹ And the Court has defined ‘hate speech’ and ‘incitement to violence’ as falling within the ‘exceptional circumstances’ justifying a custodial sentence for speech.⁴⁰⁰ Although the Court has found terms of imprisonment for hate

³⁹⁴ See, e.g., HRC, General Comment No. 34 (2011), §§34–35; ECtHR, *Tagiyev v. Azerbaijan* (App. no. 13274/08), 5 December 2019, §38; IACtHR, *Palamara Iribarne v. Chile* (Series C, no. 135), 22 November 2005, §85; ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §145.

³⁹⁵ See, e.g., ECtHR (GC), *Zana v. Turkey* (App. no. 18954/91), 25 November 1997; ECtHR, *Molnar v. Romania* (App. no. 16637/06), 23 October 2012, §25; ECtHR, *Belkacem v. Belgium* (App. no. 34367/14), 27 June 2017, §§33–37 (finding the applicant’s petition inadmissible under article 17 and forgoing review on the merits of a suspended prison sentence for incitement to discrimination, hatred and violence).

³⁹⁶ HRC, *Kim v. Republic of Korea* (Comm. no. 968/2001), 27 July 2005; HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996; see s. III.2.4.2. (Opinion). The Committee also found no violation of article 19(3) in the case of HRC, *A.K. and A.R. v. Uzbekistan* (Comm. no. 1233/2003), 31 March 2009 (where two individuals were initially convicted of speech-related offences but ultimately found guilty of association and overthrow of the constitutional order offences for recruiting and training others to undertake violent extremist activities).

³⁹⁷ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §§7–8. Cf. HRC, *Ross v. Canada* (Comm. no. 736/1997), 18 October 2000, §§11.1–11.6. See s. III.1.1. (ICCPR Article 20).

³⁹⁸ See, e.g., Human Rights Council Resolution 16/18, UN Doc. A/HRC/RES/16/18, 24 March 2011.

³⁹⁹ ECtHR, *Tagiyev v. Azerbaijan* (App. no. 13274/08), 5 December 2019, §38. The Court has also held that where fundamental rights are seriously impaired by speech, the Court considers states are ‘permitted, or even obliged, by their positive obligations under Article 8’ to regulate such expression: ECtHR, *Atamanchuk v. Russia* (App. no. 4493/11), 11 February 2020, §70. Although this language could be read as imposing mandatory penalization for such speech the case law does not establish such an obligation.

⁴⁰⁰ ECtHR, *Atamanchuk v. Russia* (App. no. 4493/11), 11 February 2020, §67. See ch. 1 (Introduction), s. II.3.1.6. (Criminal penalties for speech).

speech to be necessary and proportionate under article 10 in a limited number of cases, many of these relate to denialism, hate speech or incitement to violence cases.⁴⁰¹

For example, the Court held that a spokesperson for an extremist group could be sentenced to an 18-month suspended sentence in Belgium for uploading YouTube videos stoking ‘hatred, discrimination and violence’ against non-Muslims.⁴⁰² A six-month term of imprisonment given to a French-Romanian citizen for displaying posters containing anti-gay and anti-Roma propaganda was permissible on a threshold review.⁴⁰³ An 18-month suspended sentence was held to be proportionate when a Russian social media user uploaded a ‘mockumentary’ which expressed xenophobic views about people of non-Russian ethnic origins.⁴⁰⁴ The Court found a term of 2.5 months’ imprisonment (most of which was served on parole) proportionate where a Turkish politician expressed support for the PKK in a manner ‘likely to exacerbate an already explosive situation’ in southeast Turkey.⁴⁰⁵ And the Court upheld a suspended sentence and fines where homophobic leaflets were distributed to young people ‘who were at an impressionable and sensitive age and who had no possibility to decline to accept them.’⁴⁰⁶

In other cases, the European Court has found non-custodial criminal penalties appropriate. For instance, when a member of the Belgian parliament distributed leaflets describing non-European immigrants as criminally minded was banned from running for office for 10 years—as well as given a 10-month suspended sentence and 250 hours of community service—the Court considered that his right to free expression had not been violated and that domestic courts had shown ‘restraint’ in his sentencing.⁴⁰⁷ And when a local political commentator was fined and banned from journalistic activities for two years for authoring a demeaning article about non-Russians, the Court considered that his speech fell within the ‘exceptional circumstances’ that warranted such a punishment.⁴⁰⁸ A criminal conviction and fine of 3,000 Euros was found permissible

⁴⁰¹ See, e.g., ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012; ECtHR (GC), *Zana v. Turkey* (App. no. 18954/91), 25 November 1997; ECtHR, *Atamanchuk v. Russia* (App. no. 4493/11), 11 February 2020; ECtHR, *Soulas v. France* (App. no. 15948/03), 10 July 2008; ECtHR, *Féret v. Belgium* (App. no. 15615/07), 16 July 2009; ECtHR, *Balystė-Lideikiėnė v. Lithuania* (App. no. 72596/01), 4 November 2008; ECtHR, *Ī.A. v. Turkey* (App. no. 42571/98), 13 September 2005 (blasphemy); ECtHR, *Kilin v. Russia* (App. no. 10271/12), 11 May 2021; ECtHR, *Otto-Preminger-Institut v. Austria* (App. no. 13470/87), 20 September 1994. There are also additional relevant cases at the ECtHR which relate to other forms of speech, including terror related speech, espionage and defamation laws. See ch. 1 (Introduction), s. II.3.1.6. (Criminal penalties for speech).

⁴⁰² ECtHR, *Belkacem v. Belgium* (App. no. 34367/14), 27 June 2017, §§33–37.

⁴⁰³ ECtHR, *Molnar v. Romania* (App. no. 16637/06), 23 October 2012, §§23–26. The Court did reference proportionality, stating that ‘given the content of the posters’ the national courts’ findings were ‘relevant and sufficient’: §25.

⁴⁰⁴ ECtHR, *Kilin v. Russia* (App. no. 10271/12), 11 May 2021, §§84–95. The Court considered that domestic courts had ‘convincingly demonstrated’ a ‘clear intention to bring about the commission of related acts of hatred or intolerance’: §90.

⁴⁰⁵ ECtHR (GC), *Zana v. Turkey* (App. no. 18954/91), 25 November 1997, §60. See s. III.2.3.1. (Harm).

⁴⁰⁶ ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012, §§56–58.

⁴⁰⁷ ECtHR, *Féret v. Belgium* (App. no. 15615/07) 16 July 2009, §§34, 80. Cf. Dissenting Opinion of Judge Sajo, holding that ‘it is surprising and contrary to well-established practice that the specific criminal measures and their severity are not addressed here, even though the possibility of ten months’ imprisonment and ten years’ ineligibility (in other words, a long-term preventive interference with political speech) represents a disproportionate penalty in view of the alleged offence and the Court’s well-established jurisprudence on political speech by politicians’.

⁴⁰⁸ ECtHR, *Atamanchuk v. Russia* (App. no. 4493/11), 11 February 2020, §72. The Court considered it relevant that the commentator’s main professional activity was as an entrepreneur, and therefore that the prohibition would not have ‘significant practical consequences’. It also held the impugned statements ‘could be reasonably assessed

when a political candidate failed to delete Islamophobic comments posted by third parties on his publicly accessible Facebook ‘wall’, the Court considering that the ‘degree of liability’, ‘notoriety’ and ‘representativeness’ a local political candidate had in comparison to a private individual was relevant to its assessment of proportionality.⁴⁰⁹

The Inter-American Court of Human Rights has held that a state’s penalty for speech ‘must be proportionate to the right affected and to the responsibility of the perpetrator, so that it should be established based on the different nature and seriousness of the acts.’⁴¹⁰ It has not yet applied these principles in the context of hateful speech,⁴¹¹ but, unlike the European Court, the Inter-American Court has only approved criminal sanctions for speech in one case, considered by many to be an outlier.⁴¹²

The African Court has held that the principle of proportionality, implied by the term ‘within the law’ in article 9(2), ‘seeks to determine whether, by state action, there has been a balance between protecting the rights and freedoms of the individual and the interests of society as a whole.’⁴¹³ In making this determination, the Court considers a number of questions: ‘[a]re there sufficient reasons to justify the action? Is there a less restrictive solution? Does the action destroy the essence of the rights guaranteed by the Charter?’⁴¹⁴ Custodial sentences for violations of the laws of freedom of speech will only be lawful in ‘serious and very exceptional circumstances’ such as ‘incitement to international crimes, public incitement to hatred, discrimination or violence or threats against a person or group of people, because of specific criteria such as race, colour, religion and nationality.’⁴¹⁵ Although its practice in this area is limited, the Court has found that a 15-year prison sentence for minimizing and denying the Rwandan genocide was ‘not proportionate to the legitimate purposes’ which the sentence sought to achieve.⁴¹⁶ But in another case, the Court held that criminal offences of ‘racially motivated and xenophobic insults using a computer system’ and ‘incitement to hatred and

as stirring up base emotions or embedded prejudices in relation to the local population of non-Russian ethnicity’: ECtHR, *Atamanchuk v. Russia* (App. no. 4493/11), 11 February 2020, §64.

⁴⁰⁹ ECtHR (GC), *Sanchez v. France* (App. no. 45581/15), 15 May 2023, §201. The Court also noted that his conviction and 3,000 EUR fine (along with 1,000 EUR payment for costs) did not prevent him from being elected mayor and that ‘in an election context, the impact of racist and xenophobic discourse becomes greater and more harmful’: §§176, 208. Four judges dissented against the finding there had been no violation of article 10: Dissenting Opinion of Judge Ravarani; Dissenting Opinion of Judge Bošnjak; Joint Dissenting Opinion of Judges Wojtyczek and Zünd.

⁴¹⁰ IACtHR, *Norín Catrín et al. v. Chile* (Series C, no. 279), 29 May 2014, §374. See s. III.2.2. (Legitimacy).

⁴¹¹ IACtHR, *Norín Catrín et al. v. Chile* (Series C, no. 279), 29 May 2014, §374. See s. III.2.2. (Legitimacy).

⁴¹² IACtHR, *Mémoli v. Argentina* (Series C, no. 265), 22 August 2013, §§141, 143. The judgment has been described as a ‘regression’ by Catalina Botero-Marino, Special Rapporteur for Freedom of Expression for the Commission from 2008 to 2014: see C. Botero-Marino, ‘The Role of the Inter-American Human Rights System in the Emergence and Development of Global Norms on Freedom of Expression’, in L.C. Bollinger & A. Callamard (eds), *Regardless of Frontiers: Global Freedom of Expression in a Troubled World* (Columbia University Press 2021) 193, 194. See also E. Bertoni, ‘Setbacks and Tension in the Inter-American Court of Human Rights’ (Media Defence, 1 December 2013) (deeming the decision a ‘serious and notable setback’).

⁴¹³ ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §149, citing ACmHPR, *Zimbabwe Lawyers for Human Rights & Associated Newspapers of Zimbabwe v. Zimbabwe* (Comm. no. 284/03), 3 April 2009.

⁴¹⁴ ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §149, citing ACmHPR, *Zimbabwe Lawyers for Human Rights & Associated Newspapers of Zimbabwe v. Zimbabwe* (Comm. no. 284/03), 3 April 2009.

⁴¹⁵ ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §165.

⁴¹⁶ ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017, §162.

violence’ were ‘not disproportionate given their deterrent function’, without addressing the permissible maximum penalties under these laws.⁴¹⁷

IV. Key Divergences in International and Regional Guidance

Although limited hate speech cases have come before the Human Rights Committee, African Court and Inter-American Commission, these bodies have demonstrated that stringent conditions must be met—when it comes to intent, harm and penalties—before they will consider any punishment for hate speech to be compatible with freedom of expression. In contrast, the European Court has upheld penalties for hate speech, including criminal penalties, on a broader basis.

International bodies converge in that most cases involve balancing the right to free expression against the legitimate aim of protecting ‘the rights’ or ‘reputation’ of ‘others.’⁴¹⁸ But the European Court has departed from other bodies by finding that protecting religious sensibilities through blasphemy laws and protecting a version of history through denial laws can fall within this legitimate aim.⁴¹⁹ And where the Court considers speech to contravene an open-ended set of broad values guaranteed by the Convention it will not even inquire about whether there was a legitimate aim for the restriction.⁴²⁰

When it comes to harm, members of the Human Rights Committee have proposed that criminal penalties for hate speech offences should be limited to ‘speech that incites the commission of criminal offences or acts of violence’, the African Commission has required a ‘real likelihood and imminence of harm’ and the Inter-American Commission requires ‘an impending threat that could cause real harm.’⁴²¹ The European Court’s assessment of harm is broader, and encompasses speech that ‘could’ be ‘seen as a call for ... hatred or intolerance’⁴²² and does not need to ‘involve an explicit call for an act of violence, or other criminal acts.’⁴²³

The Human Rights Committee and other UN bodies have also placed a higher bar for the harm that is required for penalties on hate speech to be considered ‘necessary’. The Committee has held that states must demonstrate the direct and imminent nature of the harm likely to be caused by the speech, a position that has been reiterated by the

⁴¹⁷ ACtHPR, *Ajavon v. Republic of Benin* (App. no. 062/2019), 4 December 2020, §127. See s. III.2.2. (Legitimacy).

⁴¹⁸ See, e.g., HRC, *Ross v. Canada* (Comm. no. 736/1997), 18 October 2000, §11.5; *v. Sweden* (App. no. 1813/07), 9 February 2012, §49; ECtHR, *Lilliendahl v. Iceland* (App. no. 29297/18), 12 May 2020, §43; ECtHR, *Balsytė-Lideikienė v. Lithuania* (App. no. 72596/01), 4 November 2008, §73.

⁴¹⁹ See, e.g., ECtHR, *Williamson v. Germany* (App. no. 64496/17), 31 January 2019, §§16, 24; ECtHR, *Otto-Preminger-Institut v. Austria* (App. no. 13470/87), 20 September 1994, §48; ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §41.

⁴²⁰ See, e.g., ECtHR, *Norwood v. United Kingdom* (App. no. 23131/03), 16 November 2004.

⁴²¹ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §7; IACmHR, *Annual Report of the Inter-American Commission on Human Rights: Report of the Special Rapporteur for Freedom of Expression* (2009) OEA/Ser.L/V/II., Doc. 51, 397 (referring to regulation of broadcasting and not specifically to hate speech); ACmHPR, Declaration of Principles of Freedom of Expression and Access to Information in Africa (2019), Principle 23(2)(f).

⁴²² ECtHR, *Ibragimov v. Russia* (App. nos. 1413/08 and 28621/11), 28 August 2018, §§94–98.

⁴²³ *Ibid.*; ECtHR, *Stomakhin v. Russia* (App. no. 52273/07), 9 May 2018, §§129–130, see s. III.2.3.1. (Harm).

Inter-American Commission and African human rights bodies.⁴²⁴ And although the European Court will address whether speech is ‘capable’ of causing harm within its assessment of necessity, in practice, it has found restrictions to be valid without a clear demonstration of the likelihood or imminence of harm, and in its article 17 cases does not engage in this assessment at all.⁴²⁵

A final and key point of divergence has been the European Court’s willingness to consider criminal convictions, including prison terms, for hate speech to be compatible with freedom of expression.⁴²⁶ In contrast, the Human Rights Committee, Inter-American Court and African Court and Commission do not endorse criminal penalties for such speech.⁴²⁷

But ultimately, regional standards cannot be used to justify departures from international standards.⁴²⁸ There are 46 state parties to the ICCPR that are also parties to the European Convention,⁴²⁹ and where those countries provide fewer freedom of expression protections than are afforded in the ICCPR, they are in breach of their international human rights obligations even when they are complying with regional standards.⁴³⁰ This means that the more protective international standard should prevail when states are applying these standards around the globe.

V. Approach of Private Companies to Online Hate Speech

Technology and social media companies are at the forefront of responding to hate speech online: in 2020, Facebook/Meta’s removal of hate speech had risen tenfold in two years, and TikTok removed more than 100 million videos in the first half of the year.⁴³¹ A number of technology and social media companies have committed to ensuring that they address hateful speech in a manner consistent with international human rights standards.⁴³² For example, Twitter’s Rules reference freedom of expression grounded

⁴²⁴ HRC, General Comment No. 34 (2011), §35; Rabat Plan of Action, §29(f); IACmHR, *Violence against Lesbian, Gay, Bisexual, Trans and Intersex Persons in the Americas* (2015) OAS/Ser.L/V/II. Doc.36/15 Rev.2, §235; ACmHPR, Declaration of Principles of Freedom of Expression and Access to Information in Africa (2019), Principle 23(2)(f).

⁴²⁵ See, e.g., ECtHR, *Ivanov v. Russia* (App. no. 35222/04) 20 February 2007; ECtHR, *Norwood v. UK* (App. no. 23131/03) 16 November 2004. See s. III.2.3.1. (Harm).

⁴²⁶ See s. III.2.5. (Penalties). See, e.g., ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012; ECtHR (GC), *Zana v. Turkey* (App. no. 18954/91), 25 November 1997; ECtHR, *Belkacem v. Belgium* (App. no. 34367/14), 27 June 2017; ECtHR, *Molnar v. Romania* (App. no. 16637/06), 23 October 2012; ECtHR, *Garaudy v. France* (App. no. 65831/01), 24 June 2003.

⁴²⁷ With the exception of HRC, *Faurisson v. France* (Comm. no. 550/1993), 8 November 1996 (which has effectively been overruled): See s. III.2.5. (Penalties) and ACtHPR, *Ajavon v. Republic of Benin* (App. no. 062/2019), 4 December 2020.

⁴²⁸ UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §26.

⁴²⁹ See Council of Europe, ‘Chart of signatures and ratifications of Treaty 005’.

⁴³⁰ UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §26.

⁴³¹ ‘Social media’s struggle with self-censorship’ (The Economist, 22 October 2020).

⁴³² UN Special Rapporteur D. Kaye, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression* (2018) UN Doc. A/HRC/38/35, §70. See also E.M. Aswad, ‘The Future of Freedom of Expression Online’ (2018) 17 *Duke Law & Technology Review* 26, 34.

in ‘the United States Bill of Rights and the European Convention on Human Rights,’ and informed by ‘works such as United Nations Principles on Business and Human Rights.’⁴³³ Meta’s policy declares that it is ‘committed to respecting human rights as set out in ... the International Covenant on Civil and Political Rights’ and other treaties.⁴³⁴ It also has Community Standards stating that it ‘look[s] to international human rights standards to make ... judgments’ about content moderation.⁴³⁵ And the Global Network Initiative, an alliance of internet companies that includes Meta, Microsoft and other tech giants recognizes that such companies ‘have the responsibility to respect and promote the freedom of expression,’ and ‘should comply with ... internationally recognized human rights’ including the rights set out in the ICCPR,⁴³⁶ with the scope of article 19(3) to be ‘read within the context of further interpretations issued by international human rights bodies, including the Human Rights Committee and the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression.’⁴³⁷ The decisions of Meta’s Oversight Board—a group of experts established in 2018 to determine ‘what to take down, what to leave up and why’⁴³⁸—also assess whether Meta’s conduct complies with ‘international human rights standards.’⁴³⁹

For example, the Oversight Board considered a post by a Facebook user in Myanmar saying that there was something wrong with Muslim men psychologically and questioning the lack of response by Muslims to the treatment of Uyghurs. Facebook removed this on the basis that it was unprotected hate speech—defined as a ‘direct attack against people’ on the basis of ‘protected characteristics.’⁴⁴⁰ But the Oversight Board disagreed, holding that ‘restoring the post is consistent with international human rights standards,’ including article 19 of the ICCPR, General Comment No. 34, the Rabat Plan of Action and the UN Special Rapporteur’s report on Online Hate Speech.⁴⁴¹ Applying these principles, the Board found that the content did not constitute ‘advocacy of

⁴³³ Twitter (known as X since July 2023), ‘Defending and respecting the rights of people using our service’; See also V. Türk, ‘Open letter from Volker Türk, United Nations High Commissioner for Human Rights, to Mr. Elon Musk, Chief Executive Officer at Twitter’ (5 November 2022).

⁴³⁴ Meta, ‘Corporate Human Rights Policy’.

⁴³⁵ Meta, ‘Facebook Community Standards’. See also TikTok, ‘Upholding human rights’ (‘Our philosophy is informed by the International Bill of Human Rights ... and the United Nations Guiding Principles on Business and Human Rights’); Facebook Newsroom, ‘Hard Questions: Where Do We Draw the Line on Free Expression?’ (9 August 2018) (‘We look for guidance in documents like Article 19 of the International Covenant on Civil and Political Rights (ICCPR), which set standards for when it’s appropriate to place restrictions on freedom of expression’).

⁴³⁶ Global Network Initiative, ‘GNI Principles on Freedom of Expression and Privacy’.

⁴³⁷ *Ibid.*, n 7. The GNI Principles have also ‘been drafted with reference to’ the Johannesburg Principles: *Ibid.*, n 9. And Facebook, Google, Microsoft and Twitter also have policies in place to engage with human rights experts and civil society organizations to ensure they are correctly implementing these standards. See, e.g., Meta, ‘Facebook Community Standards’.

⁴³⁸ Oversight Board, ‘Oversight Board’.

⁴³⁹ Oversight Board, ‘Case decision 2021-015-FB-UA Asking for Adderall’; Oversight Board, ‘Case decisions and policy advisory opinions’. See also Oversight Board, ‘Case decision 2020-003-FB-UA Armenians in Azerbaijan’ (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’).

⁴⁴⁰ Meta, ‘Community Standards: Hate Speech’.

⁴⁴¹ Oversight Board, ‘Case decision 2020-002-FB-UA Myanmar post about Muslims’. The decisions of the Oversight Board are ‘binding,’ however, if a decision or policy advisory opinion includes recommendations, Meta is only required to analyse its operational procedures and consider such recommendations: see Oversight Board, ‘Oversight Board Charter, 7–8.’

religious hatred constituting incitement to discrimination, hostility or violence, which states are required to prohibit under ICCPR Article 20, para. 2'. The Board then considered 'whether this content could be restricted under ICCPR Article 19, para 3', and concluded that its removal was not necessary to protect the rights of others.⁴⁴²

Similarly, in a case approving Facebook's decision to remove a post which used a slur to insult Azerbaijanis, the Board noted that 'Facebook has ... indicated that it looks to authorities such as the ICCPR and the Rabat Plan of Action when making content decisions'. Applying these, it found that Facebook's removal of the post met 'the three-part test of legality, legitimacy and necessity and proportionality' required by article 19, particularly since the speech was posted during the fighting in Nagorno-Karabakh, and the dangers of dehumanizing slurs proliferating in a way that escalates into violence.⁴⁴³

The Oversight Board recently assessed Meta's balancing act between limiting the spread of violent hate speech and protecting valid forms of political expression in a case regarding a Facebook post calling for the end of Iranian leader Ayatollah Khamenei's regime.⁴⁴⁴ In a public group describing itself as supporting freedom in Iran, a Facebook user posted in July 2022 a caricature of Iran's Supreme Leader with a text bubble saying that being a woman is forbidden, and a caption calling for death to the 'anti-women Islamic government' and its 'filthy leader Khamenei', as well as encouraging Iranian women not to collaborate with the oppression of women. This was posted just months before the death of Mahsa Amini in custody which sparked widespread protests in Iran.⁴⁴⁵ Meta removed the post and restricted the user's Facebook access on the basis that it prohibits calls for death or high-severity violence, but when the Board selected the case for review, Meta reversed its decision on the basis that the post fell within its 'newsworthiness' exception, where keeping content online is 'in the public interest'.

The Oversight Board found that removing the post was inconsistent with Meta's 'human rights responsibilities as a business' because the post was a 'rhetorical, political slogan, not a credible threat' and 'posed very little risk of inciting violence'.⁴⁴⁶ However, the Board did not consider all 'death to' statements to be the same: for example, 'death to' statements against Salman Rushdie would 'pose a much more significant risk' in light of the fatwa against him and ongoing concerns for his safety. And 'death to' statements

⁴⁴² Oversight Board, 'Case decision 2020-002-FB-UA Myanmar post about Muslims'.

⁴⁴³ Oversight Board, 'Case decision 2020-003-FB-UA Armenians in Azerbaijan'.

⁴⁴⁴ Oversight Board, 'Case decision 2022-013-FB-UA Iran protest slogan'.

⁴⁴⁵ F. Fassihi & C. Engelbrecht, 'Tens of Thousands in Iran Mourn Mahsa Amini, Whose Death Set Off Protests' (The New York Times, 26 October 2022).

⁴⁴⁶ Oversight Board, 'Case decision 2022-013-FB-UA Iran protest slogan'. Because the Board found that the post did not violate Meta's Violence and Incitement Community Standard, it considered that an assessment of the 'newsworthiness allowance was not required', which is the balancing test that Meta conducts comparing the public interest in the content against the risk of harm. Notwithstanding this finding, the Board held that when Meta applied the newsworthiness allowance to the post 'it should have been scaled to apply to all "marg bar Khamenei" slogans, regardless of the speaker'.

during the events of the January 6 riots in Washington, D.C., were not comparable, as ‘politicians were clearly at risk and ‘death to’ statements are not generally used as political rhetoric in English.’⁴⁴⁷ The Board also considered the political context in Iran, noting that in an environment where ‘the Iranian government systematically represses freedom of expression . . . it is vital that Meta supports users’ voice.’ The Board ultimately recommended that Meta amend its Violence and Incitement Community Standard to provide ‘the criteria used to determine when rhetorical threats against heads of state are permitted’, and that these criteria should ‘protect clearly rhetorical political speech, used in protest contexts, that does not incite violence, and should take language and context into account.’⁴⁴⁸

Although some internet companies have committed themselves to the minimum international standards on speech, it is not clear how they will address divergences in the guidance. Twitter grounds its values in both the First Amendment *and* the European Convention, two very different bodies of law.⁴⁴⁹ The Oversight Board, on the other hand, has stated that it would not ‘apply the First Amendment of the US Constitution’ as this ‘does not govern the conduct of private companies’, but also noted that ‘in many relevant respects the principles of freedom of expression reflected in the First Amendment are similar or analogous to the principles of freedom of expression in ICCPR Article 19’.⁴⁵⁰ And the Board has recognized that human rights standards designed for countries may need to be transposed or interpreted differently when applied to companies under some circumstances.⁴⁵¹

VI. Recommendations

The following recommendations are based on international standards applicable to hate speech, or where there is a divergence or lacunae in such standards, a suggestion as to the best policy or approach. They are addressed to states because states have ratified the international treaties that underlie the standards. However, they are also intended to guide private companies seeking to apply international standards in line with the UN Guiding Principles on Business and Human Rights.⁴⁵²

⁴⁴⁷ Oversight Board, ‘Case decision 2022-013-FB-UA Iran protest slogan.’

⁴⁴⁸ Ibid.

⁴⁴⁹ See ss III.2. (Discretionary Restrictions on Hate Speech) and II.2.3.2. (United States).

⁴⁵⁰ Oversight Board, ‘Case decision 2021-001-FB-FBR Former President Trump’s suspension’, 28.

⁴⁵¹ The Board has noted with approval the fact that the UN Special Rapporteur sees greater leeway for speech restriction by a company than by a government. See Oversight Board, ‘Case decision 2020-003-FB-UA Armenians in Azerbaijan’. See also Oversight Board, ‘Case decision 2021-001-FB-FBR Former President Trump’s suspension’, 25.

⁴⁵² See Twitter (known as X since July 2023), ‘Defending and respecting the rights of people using our service’; Meta, ‘Corporate Human Rights Policy’; see s. V. (Approach of Private Companies to Online Hate Speech). As all restrictions to speech, discretionary or purportedly mandatory, must comply with article 19(3) of the ICCPR, our recommendations regarding the limits of discretionary restrictions to speech apply with equal force to mandatory restrictions.

1. Blasphemy laws are not compatible with freedom of expression under international human rights law and should be repealed.

Restrictions on speech cannot be vague or overly broad, as individuals must know how to regulate their conduct to avoid penalties imposed through the law.⁴⁵³ And to be legitimate, hate speech laws must fall within the two exhaustive aims of article 19(3): respecting the rights or reputations of others, or the protection of national security, public order or public health or morals.⁴⁵⁴

Although at least 95 countries still punish blasphemy with criminal penalties,⁴⁵⁵ blasphemy laws do not meet this test.⁴⁵⁶

The UN Human Rights Committee has made clear that ‘[p]rohibitions of displays of lack of respect for a religion or other belief system, including blasphemy laws, are incompatible with’ article 19 of the Covenant, and that ‘criticism of even the most deeply-held convictions of the adherents of a religion is protected by freedom of expression.’⁴⁵⁷ Similarly, the Inter-American Court has not tolerated censorship of speech simply on the basis that it offends religious sensibilities.⁴⁵⁸ And some European bodies, including the European Union and the Venice Commission (an expert body that forms part of the Council of Europe), have followed this approach and recommended the decriminalization of blasphemy laws.⁴⁵⁹

This approach is preferable to the one adopted by the European Court of Human Rights which has, in effect, allowed restrictions to speech on the basis of distasteful content—rather than based on the intent of the speaker and the harm caused by the speech—even in cases involving criminal penalties. Viewpoint-based restrictions set a worrying precedent whereby the state decides which content is or is not acceptable or even whether the

⁴⁵³ HRC, General Comment No. 34 (2011), §25; UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §§6(a), 32, 46; ECtHR, *Dink v. Turkey* (App. nos. 2668/07 & others), 14 September 2020, §114 (although cf. s. III.2.1. (Legality), discussing the Court’s more lenient application of this principle); IACtHR, *Usón Ramírez v. Venezuela* (Series C, no. 207), 20 November 2009, §§5; ACtHPR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §131.

⁴⁵⁴ ICCPR Art.19(3). See ss III.2.2. (Legitimacy) and III.2 (Discretionary Restrictions on Hate Speech).

⁴⁵⁵ See United States Commission on International Religious Freedom, ‘Legislation Factsheet: Blasphemy (September 2023)’; see ss II.1.1. (Type of speech) (Case Study: Blasphemy Laws) and II.1.1. (Type of speech) (Case Study: Denial laws).

⁴⁵⁶ HRC, General Comment No. 34 (2011), §48. See also HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, §10.5; Venice Commission, *Report on the Relationship between Freedom of Expression and Freedom of Religion: The Issue of Regulation and Prosecution of Blasphemy, Religious Insult and Incitement to Religious Hatred* (2008) CDL-AD(2008)026, §92; EU Guidelines on the promotion and protection of freedom of religion or belief, §32; OSCE, ‘OSCE Representative on Freedom of the Media welcomes Irish referendum on blasphemy’ (2 October 2017). See s. III.2.2. (Legitimacy) (Case Study: International Standards on Blasphemy Laws).

⁴⁵⁷ HRC, General Comment No. 34 (2011), §48; HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, §10.5. See also HRC, *Concluding Observations, Kuwait* (2016) UN Doc. CCPR/C/KWT/CO/3, §§39, 41; HRC, *Concluding Observations, Russian Federation* (2015) UN Doc. CCPR/C/RUS/CO/7, §19. See also e.g., WGAD, *Mejri v. Tunisia* (Opinion no. 29/2013), 30 August 2013 §18; WGAD, *Amer v. Egypt* (Opinion no. 35/2008), 20 November 2008, §38.

⁴⁵⁸ See IACtHR, ‘*The Last Temptation of Christ*’ (*Olmedo-Bustos et al.*) v. *Chile* (Series C, no. 73), 5 February 2001, §§71–73, 76–80 (holding that Chile’s prohibition of the exhibition of the film ‘The Last Temptation of Christ’ on the basis that it ‘deformed and diminished’ the image of Christ constituted prior censorship in violation of article 13, and did not violate article 12 (freedom of religion)).

⁴⁵⁹ Venice Commission, *Report on the Relationship between Freedom of Expression and Freedom of Religion: The Issue of Regulation and Prosecution of Blasphemy, Religious Insult and Incitement to Religious Hatred* (2008) CDL-AD(2008)026, §92; EU Guidelines on the promotion and protection of freedom of religion or belief; OSCE, ‘OSCE Representative on Freedom of the Media welcomes Irish referendum on blasphemy’ (2 October 2017).

content has been expressed in an acceptable way. The decision in *E.S. v. Austria*, upholding a conviction under Austria's blasphemy law for referring to Prophet Muhammad's marriage to a six-year-old girl as paedophilia epitomizes the low bar for allowing a conviction for blasphemous speech in Europe.⁴⁶⁰ Although such speech may be offensive or even reprehensible, the European Court's approval of criminal penalties in such instances runs counter to the minimum international standards that bind all states.

As the UN Special Rapporteur has said, 'article 19 protects *individuals* ... [it does not] protect *ideas or beliefs* from ridicule, abuse, criticism or other "attacks" seen as offensive.'⁴⁶¹ And as the UN-approved Rabat Plan of Action puts it, blasphemy laws are 'counterproductive, since they may result in de facto censure of all inter-religious or belief and intra-religious or belief dialogue, debate and criticism, most of which could be constructive, healthy and needed.'⁴⁶² We therefore urge states to repeal such provisions and in doing so ensure compliance with international law.

Repealing blasphemy provisions does not of course mean that any attack against religious groups or on the grounds of religion should go unaddressed. There are other ways to combat such speech without resorting to the blunt instrument of the criminal law. As recognized by the Human Rights Council, the 'open public debate of ideas, as well as interfaith and intercultural dialogue ... can be among the best protections against religious intolerance.'⁴⁶³ The fact that blasphemous speech may lead to a violent response from those who are incensed by the speech does not change this analysis: international human rights law does not allow for a 'heckler's veto,' which would mandate the stifling of speakers when those who are offended choose to show their displeasure through harmful acts.'⁴⁶⁴ If the speech constitutes not just blasphemy but incitement to discrimination, hostility or violence, international law provides that it should be sanctioned.⁴⁶⁵

2. 'Denial laws' are not compatible with international human rights law and should be repealed.

The UN Human Rights Committee has held that opinions denying or calling into question genocide⁴⁶⁶ or historical narratives cannot be penalized.⁴⁶⁷ We endorse the Committee's position that '[l]aws that penalize the expression of opinions about historical facts are incompatible with the obligations that the [ICCPR] imposes on state parties in relation to the respect for freedom of opinion and expression.'⁴⁶⁸ And we agree

⁴⁶⁰ ECtHR, *E.S. v. Austria* (App. no. 38450/12), 25 October 2018, §§43–57.

⁴⁶¹ UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §21 (emphasis added).

⁴⁶² Rabat Plan of Action, §19.

⁴⁶³ Human Rights Council Resolution 16/18, UN Doc. A/HRC/RES/16/18, 24 March 2011, §4.

⁴⁶⁴ UN Special Rapporteur D. Kaye, *Promotion and protection of the right to freedom of opinion and expression* (2019) UN Doc. A/74/486, §10; E. Aswad, 'To Ban or Not to Ban Blasphemous Videos' (2013) 44 *Georgetown Journal of International Law* 1313, 1322.

⁴⁶⁵ See ICCPR Art. 20; CERD Art. 4.

⁴⁶⁶ Professor Irwin Cotler endorses all recommendations in this book except for this one.

⁴⁶⁷ HRC, General Comment No. 34 (2011), §49; CERD Committee, General Recommendation No. 35 (2013), §14.

⁴⁶⁸ HRC, General Comment No. 34 (2011), §49.

with the Committee that international law ‘does not permit general prohibition of expressions of an erroneous opinion or an incorrect interpretation of past events.’⁴⁶⁹

Yet the European Court and a number of European nations have adopted the opposite position,⁴⁷⁰ with the European Court finding that denial of the Holocaust is unprotected under any circumstances and that a claim that such speech is protected cannot even be considered by the Court.⁴⁷¹ In such decisions the European Court has allowed criminal penalties to be meted out against speech on the basis of its content, rather than the intent of the speaker and likely harm caused by the speech, a notion that sets a worrying precedent whereby a state or court decides which viewpoint is or is not acceptable or even whether a viewpoint has been expressed in an acceptable way.⁴⁷² But ultimately the intent of the speaker, the harm that is caused by the speech, and any extenuating or aggravating circumstances are best assessed on a case-by-case basis.⁴⁷³

The European Court’s position that denial of the Holocaust is criminal but denial of the Armenian genocide is acceptable is also problematic as it creates a ‘speculative hierarchy between the Holocaust and other twentieth-century mass atrocities.’⁴⁷⁴ It also begs many questions: is it only historic events that cannot be denied? Does it have to be a past genocide or can it be an ongoing one? Does there need to be a court judgment confirming such a genocide—if so, which court counts? Does this mean that denial of crimes against humanity, or torture, or rape can also be criminalized? Why these and not other viewpoints? Instead of opening floodgates or engaging in unprincipled line-drawing, statements of opinion should be protected unless they violate article 20 of the ICCPR⁴⁷⁵ and denial laws should be abolished.

3. States should only criminalize hate speech if it is intended to incite the commission of serious criminal offences, acts of violence or similar action and there is at least a likelihood, if not probability, that such offences or acts will imminently occur as a direct result of the speech.

States can only impose penalties on hate speech to the extent that the penalty is necessary and proportionate to advancing a legitimate aim.⁴⁷⁶ Although the vast majority

⁴⁶⁹ Ibid.

⁴⁷⁰ In 2020, Facebook also announced that it would prohibit ‘any content that denies or distorts the Holocaust’: Meta, ‘Removing Holocaust Denial Content’ (12 October 2020).

⁴⁷¹ See s. III.2. (Discretionary Restrictions on Hate Speech). The ACtHPR has also expressed openness to laws which penalize negation or minimization of genocide in the context of the Rwandan genocide, presenting a rare instance in which the ACtHPR has veered away from HRC jurisprudence. See ACtHPR, *Umuhoza v. Rwanda* (App. no. 003/2014), 24 November 2017, §158; see s. III.2.4.2. (Opinion). See also H. Cannie & D. Voorhoof, ‘The Abuse Clause and Freedom of Expression in the European Human Rights Convention: An Added Value for Democracy and Human Rights Protection?’ (2011) 29 *Netherlands Quarterly of Human Rights* 54, 67–68.

⁴⁷² See ECtHR, *Garaudy v. France* (App. no. 65831/01), 24 June 2003, §1.

⁴⁷³ See s. III.2.2. (Legitimacy).

⁴⁷⁴ A. Cherviatsova, ‘Memory as a Battlefield: European Memorial Laws and Freedom of Speech’ (2021) 25 *The International Journal of Human Rights* 675, 684.

⁴⁷⁵ But see s. VI.3.2 (Recommendations).

⁴⁷⁶ See s. III.2.3. (Necessity).

of state hate speech laws are criminal in nature,⁴⁷⁷ international standards dictate that criminal penalties for speech should only be resorted to in exceptional circumstances.⁴⁷⁸ The lack of cases condoning criminal penalties for hate speech by international and regional human rights bodies is a clear marker that hate speech will rarely be considered sufficiently exceptional to warrant criminal sanctions, especially if these include imprisonment.⁴⁷⁹ Even when it comes to mandatory penalties for speech imposed by treaty, international guidance provides that states should use criminal sanctions ‘only [in] serious and extreme instances.’⁴⁸⁰ And such standards cannot be read to require that a prosecution—let alone a conviction—must eventuate.⁴⁸¹

3.1. The speaker must have ‘intended’ to incite harm.

The Genocide Convention has an explicit and high intent requirement: the intent to directly and publicly incite others to commit genocide.⁴⁸² Although article 20 of the ICCPR does not contain an express intent requirement, persuasive soft law guidance construes ‘advocacy’ as ‘the explicit, *intentional*, public and active support and promotion of hatred towards the target group’.⁴⁸³ General Comment No. 35 holds that states should take into account ‘the intention of the speaker’ when applying article 4 of CERD, and must demonstrate that the speaker ‘seeks to influence others to engage in certain forms of conduct . . . through advocacy or threats.’⁴⁸⁴ Finally, the UN-approved Rabat Plan of Action includes intent as one of the six factors that states should consider when criminalizing hateful speech and holds that ‘[n]egligence and recklessness are not sufficient for [a speech] act to be an offence.’⁴⁸⁵ Strict liability offences or laws that do

⁴⁷⁷ See Alkiviadou, Mchangama & Mendiratta (n 19); see s. III.2.5. (Penalties).

⁴⁷⁸ See, e.g., IACtHR, *Álvarez Ramos v. Venezuela* (Series C, no. 380), 30 August 2019, §§119, 120; ACtHR, *Konaté v. Burkina Faso* (App. no. 004/2013), 5 December 2014, §165; HRC, General Comment No. 34 (2011), §47 (in the context of defamation); CERD Committee, General Recommendation No. 35 (2013), §12.

⁴⁷⁹ Even the European Court has only upheld criminal sanctions for hate speech in a small fraction of its decisions, even less so where the speaker was sentenced to a term of imprisonment. See, e.g., ECtHR, *Atamanchuk v. Russia* (App. no. 4493/11), 11 February 2020; ECtHR, *Soulas v. France* (App. no. 15948/03), 10 July 2008; ECtHR, *Ivanov v. Russia* (App. no. 35222/04), 20 February 2007; ECtHR, *Nix v. Germany* (App. no. 35285/16), 13 March 2018; ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07), 9 February 2012; ECtHR, *Belkacem v. Belgium* (App. no. 34367/14), 27 June 2017; ECtHR, *Féret v. Belgium* (App. no. 15615/07), 16 July 2009; ECtHR, *Le Pen v. France* (App. no. 18788/09), 7 May 2010; ECtHR, *Molnar v. Romania* (App. no. 16637/06), 23 October 2012; ECtHR, *Norwood v. United Kingdom* (App. no. 23131/03), 16 November 2004.

⁴⁸⁰ UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §47. See IACmHR, *Violence against Lesbian, Gay, Bisexual, Trans and Intersex Persons in the Americas* (2015) OAS/Ser.L/V/II. Doc.36/15 Rev.2, ch. IV, s. F. (proposing non-legal measures to counter hate speech, including preventive and educational mechanisms).

⁴⁸¹ HRC, *Rabbae v. The Netherlands* (Comm. no. 2141/2011), 14 July 2016, §10.7 (‘The obligation under article 20 (2), however, does not extend to an obligation for the State party to ensure that a person who is charged with incitement to discrimination, hostility or violence will invariably be convicted by an independent and impartial court of law’). The CERD Committee has, however, also held that ‘effective implementation’ of article 4 is ‘characteristically achieved through investigations of offences set out in the Convention and, where appropriate, the prosecution of offenders’. See CERD Committee, General Recommendation No. 35 (2013), §17 (emphasis added); see also §12.

⁴⁸² See s. III.1.3. (Genocide Convention Article 3).

⁴⁸³ UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §§43–44 (emphasis added); Rabat Plan of Action, §21, n 5.

⁴⁸⁴ CERD Committee, General Recommendation No. 35 (2013), §16.

⁴⁸⁵ Rabat Plan of Action, §29(c). See also UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §50. See also IACmHR, *Violence against Lesbian, Gay, Bisexual, Trans and Intersex Persons in the Americas* (2015) OAS/Ser.L/V/II. Doc.36/15 Rev.2, §235; ACmHPR, Declaration of Principles of Freedom of Expression and Access to Information in Africa (2019), Principle 23(2)(c).

not set out a sufficiently stringent intent requirement are therefore at odds with minimum international standards. Recklessness could, however, be sufficient to render the speaker civilly liable where speech is restricted on a discretionary basis under article 19 of the ICCPR.⁴⁸⁶

3.2. The harm that the speaker intended to incite should be ‘the commission of criminal offences, acts of violence or similar action.’

A discrepancy exists between international instruments as to what harm justifies the penalization or criminalization of hate speech. States’ obligation to restrict speech under article 20 of the ICCPR and article 4 of CERD only applies to speech that amounts to ‘incitement’⁴⁸⁷ but the incited harms that can justify penalization are broad: including ‘hostility’ under the ICCPR, ‘contempt’ and ‘hatred’ under CERD as well as the common harms of ‘discrimination’ and ‘violence.’⁴⁸⁸ A higher harm standard is required by the American Convention: ‘lawless violence’ or ‘similar action,’ similar to US First Amendment principles.⁴⁸⁹ And a similar standard was articulated by members of the Human Rights Committee who held that speech should incite to ‘criminal offences or acts of violence’—at least before mandatory penalties or criminal sanctions can be justified.⁴⁹⁰

This higher standard is preferable as it avoids overbroad criminalization. As noted by members of the Human Rights Committee, international law ‘urges great caution in the imposition of criminal penalties that punish speech.’⁴⁹¹ And terms such as ‘hostility,’ ‘hatred’ and ‘contempt’ are subjective states of mind rather than objective hate-fuelled acts that result from the speech.⁴⁹² Authoritarian regimes can easily manipulate such terms to silence criticism and would no doubt be glad to say that they were just following international agreed standards in doing so.⁴⁹³ Such vague concepts, if used in criminal laws, may well also violate the principle of legality (also known as *nullum crimen sine lege*) as it would be difficult for an individual to predict

⁴⁸⁶ See ch. 2 (Insulting Speech), s. IV. (Recommendations) and ch. 4 (False Speech), s. IV. (Recommendations).

⁴⁸⁷ With incitement incorporating an ‘imminent risk’ of harm: see HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §4; UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §43; Rabat Plan of Action, §21, n 5; CERD Committee, General Recommendation No. 35 (2013), §13(b), (d).

⁴⁸⁸ CERD Committee, General Recommendation No. 35 (2013), §13(b), (d).

⁴⁸⁹ See s. III.1.4. (Mandatory restrictions at the regional level).

⁴⁹⁰ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §7. See s. III.1.1. (ICCPR Article 20).

⁴⁹¹ HRC, *Rabbae v. The Netherlands* (Comm. no. 2124/2011), 14 July 2016, Individual Concurring Opinion of Committee members Sarah Cleveland and Mauro Politi, §7.

⁴⁹² Some UN documents have suggested that terms such as ‘hostility’ be interpreted to mean ‘a manifestation of hatred beyond a mere state of mind’ but this has not been adopted by international courts and still leaves terms such as ‘hatred’ and ‘contempt’ undefined: See, e.g., UN Special Rapporteur F. La Rue, *Promotion and protection of the right to freedom of opinion and expression* (2012) UN Doc. A/67/357, §44(e).

⁴⁹³ See s. III.2.5. (Penalties).

what speech may lead to a criminal penalty.⁴⁹⁴ And they are overbroad: as polarized political views are evident in much of today's news media, much of the content of our newspapers or news programs might easily pass the test of inspiring 'hostility' or 'contempt' in audiences.

States should therefore consider entering a reservation to article 20 of the ICCPR and article 4 of CERD in a manner that would allow them to apply these provisions as imposing permission but not a requirement to penalize speech, and to apply a high intent and harm requirement to any penalization—such as an intent to cause violence or illegal acts that may cause serious physical injury or death through speech—rather than the more vague ills such as hostility, contempt and discrimination that are mentioned in these treaty provisions.

3.3. There should be a 'probability' that the harm the speaker intended to invite would 'imminently occur as a direct result of the speech'.

The Human Rights Committee has held that states must 'demonstrate in specific and individualized fashion the precise nature of the threat' posed by speech and establish a 'direct and immediate connection between the expression and the threat' if it is to be penalized.⁴⁹⁵ Under the UN-approved Rabat Plan of Action, this means that there should be 'a *reasonable probability* that the speech would succeed in *inciting actual action* against the target group, recognizing that such causation should be rather direct.'⁴⁹⁶ The CERD Committee has provided a similar threshold, being that (at least in criminal cases) states must demonstrate the 'imminent risk of likelihood that the conduct . . . intended by the speaker will result from the speech.'⁴⁹⁷

Yet currently, in practice, speech is often sanctioned without any demonstrated connection between speech and harm, let alone an objective probability (meaning that it is more probable than not) of harm that is imminent.⁴⁹⁸ Weak causation tests such as the European Court's standard of whether speech 'could' be 'seen as a call for violence, hatred or intolerance',⁴⁹⁹ or total disregard of causation by virtue of article 17,⁵⁰⁰ leave too much speech unprotected or subject to viewpoint-based restrictions.⁵⁰¹

⁴⁹⁴ See ch. 4 (False Speech), s. III.1. (Legality).

⁴⁹⁵ HRC, General Comment No. 34 (2011); Rabat Plan of Action, §29(f); UN Special Rapporteur on the right to freedom of opinion and expression, Frank LaRue, *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression (4 September 2013)*, UN Doc. A/68/362, §§52, 53.

⁴⁹⁶ Rabat Plan of Action, §29(f) (emphasis added).

⁴⁹⁷ CERD Committee, General Recommendation No. 35 (2013), §16.

⁴⁹⁸ See s. III.2.3.1. (Harm). See, e.g., ECtHR (GC), *Zana v. Turkey* (App. no. 18954/91, 25 November 1997, §60; ECtHR, *Molnar v. Romania* (App. no. 16637/06), 23 October 2012, §23.

⁴⁹⁹ ECtHR, *Ibragimov v. Russia* (App. nos. 1413/08 & 28621/11), 28 August 2018, §98; ECtHR (GC), *Perinçek v. Switzerland* (App. no. 27510/08), 15 October 2015, §207.

⁵⁰⁰ See, e.g., ECtHR, *Lilliendahl v. Iceland* (App. no. 29297/18), 12 May 2020, §36.

⁵⁰¹ For example, blasphemy or denialism laws which prohibit speech on the basis of content rather than harm or intent: See s. VI. (Recommendations).

-
4. When it is legitimate to penalize hate speech, this can be on the ground of prohibited harm to individuals on the basis of sexual orientation, gender, political affiliation, or other protected characteristics, not just discrimination based on race, ethnicity or religion.

States' obligations to limit hate speech vary based on the target of the speech: article 20 of the ICCPR applies to 'national, racial or religious' hatred. Article 4 of CERD only applies to 'racial' hatred, though the CERD Committee has construed this to extend to 'colour, descent or national or ethnic origin'.⁵⁰²

The Inter-American approach covers characteristics other than nationality, race and religion and extends to speech that is hateful on the basis of 'sexual orientation, gender identity, and bodily diversity'.⁵⁰³ Similarly, the European Court has held that 'discrimination based on sexual orientation is as serious' as discrimination based on 'race, origin or colour'.⁵⁰⁴

International human rights treaties were drafted in the wake of World War II with a specific threat of totalitarianism in mind. But it is time to recognize that the protection in international treaties for the 'rights of others' includes other characteristics such as those outlined in article 26 of the ICCPR, which prohibits 'discrimination on any ground such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status'. But balancing the protection from discrimination against the right to free speech would need to remain in line with the other recommendations related to strict requirements of intent, causation, harm and guidance of proportionality of penalties as set out above.

5. Tech companies should recognize international human rights standards as a floor, not a ceiling, of free speech protection.

Social media and technology companies should uphold the commitment to regulating speech in line with the GNI principles, which recognize that international human

⁵⁰² See CERD Committee, General Recommendation No. 35 (2013), §13 ('States parties are required ... [to] sanction as offences punishable by law: ... All dissemination of ideas based on racial or ethnic superiority or hatred ... Incitement to hatred, contempt or discrimination against members of a group on grounds of their *race, colour, descent, or national or ethnic origin*') (emphasis added); §6 ('article 4 ... *forbids discrimination on grounds of race, colour, descent, or national or ethnic origin*—such as indigenous peoples, descent-based groups, and immigrants or non-citizens, including migrant domestic workers, refugees and asylum seekers, as well as speech directed against women members of these and other vulnerable groups') (emphasis added).

⁵⁰³ IACmHR, *Violence against Lesbian, Gay, Bisexual, Trans and Intersex Persons in the Americas* (2015) OAS/Ser.L/V/II. Doc.36/15 Rev.2, §13; see s. III.1.4. (Mandatory restrictions at the regional level).

⁵⁰⁴ ECtHR, *Vejdeland v. Sweden* (App. no. 1813/07, 9 February 2012, §55).

rights law including article 19 of the ICCPR should be respected and ‘read within the context of further interpretations issued by international human rights bodies.’⁵⁰⁵ International human rights standards are, however, a floor rather than a ceiling of protection for speech—and should be treated as such by private companies that have a role in regulating speech.

⁵⁰⁵ Global Network Initiative, ‘GNI Principles on Freedom of Expression and Privacy’, n 7. See also Forum of Information & Democracy, ‘Principles on Information & Democracy’ (endorsed by 51 states from across the world that comprise the International Partnership for Information and Democracy).



Human Rights
Institute