PROTECTING BOTH SIDES OF THE CONVERSATION: TOWARDS A CLEAR INTERNATIONAL STANDARD FOR HATE SPEECH REGULATION

MEERA CHANDRAMOULI*

1. INTRODUCTION

The First Amendment to the United States Constitution represents a wholehearted commitment to an individual's right to free expression.[1] Historically, the Supreme Court has permitted the abridgment of that right only in very particular circumstances. In the majority of its decisions, the Court has adopted a speech-protective stance, prioritizing an individual's contribution to the common marketplace of ideas over government intervention.[2] The notion is that the addition of any idea, even if it is disfavored, outlandish, or small-minded, to society helps to preserve an atmosphere of robust public debate.[3]

The international community, in contrast, is warier of the externalities that accompany unchecked free expression and has, as

---

* Comments Editor, *University of Pennsylvania Journal of International Law*, Vol. 34. J.D. Candidate, 2013, University of Pennsylvania Law School; B.A., 2009, New York University. I would like to thank the editors of the *University of Pennsylvania Journal of International Law* for their tireless efforts. I would also like to thank my family and friends for their support, humor, and wisdom. All errors are my own.

1 U.S. CONST. amend. I ("Congress shall make no law . . . abridging the freedom of speech . . . .").

2 *See* Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting) ("[W]hen men have realized that time has upset many fighting faiths, they may come to believe . . . that the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market . . . .").

3 *See, e.g.*, Snyder v. Phelps, 131 S. Ct. 1207, 1219 (2011) (upholding the constitutionality of picketing at a soldier's funeral and stressing that "[public] speech cannot be restricted simply because it is upsetting or arouses contempt"); Texas v. Johnson, 491 U.S. 397, 414 (1989) (relying on the "bedrock principle underlying the First Amendment . . . that the government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable" to invalidate a flag burning statute); Gertz v. Robert Welch, Inc., 418 U.S. 323, 339 (1974) ("Under the First Amendment there is no such thing as a false idea.").

a result, moved towards broader regulation of hate speech,[4] or "speech designed to promote hatred on the basis of race, religion, ethnicity or national origin."[5] The reason for this movement stems in part from a post-World War II notion that characterizes hate speech not merely as the expression of a viewpoint, but also as an instrument of subjugation.[6] In its desire to equally uphold the tenets of free expression and freedom from discrimination, however, the international community has failed to produce an articulate, comprehensive standard for hate speech regulation.

In this Comment, I will suggest that incorporating frameworks from U.S. First Amendment jurisprudence into international law may enhance the precision of the current international standard for hate speech regulation. In Part 2, I will examine the protections offered to hate speech in First Amendment jurisprudence in the United States. In order to understand the American approach to regulating hate speech, it is necessary to first examine two related categorical exceptions to First Amendment protection into which hate speech may fall: words that incite violence and "fighting words."[7] I will analyze the Court's treatment of these exceptions in

[4] In this comment, I will use the term "hate speech" to denote not only speech, but also expressive conduct, *see, e.g.*, *Johnson*, 491 U.S. 397 (holding that the act of burning the American flag in protest constitutes expressive conduct); U.S. v. O'Brien, 391 U.S. 367, 376 (1968) (addressing whether the act of burning draft cards can be "protected 'symbolic speech' within the First Amendment"), and advocacy, *see* Brandenburg v. Ohio, 395 U.S. 444 (1969) (outlining the current standard by which the U.S. Supreme Court evaluates governmental regulation of advocacy that directly incites violence).

[5] Michel Rosenfeld, *Hate Speech in Constitutional Jurisprudence: A Comparative Analysis,* 24 CARDOZO L. REV. 1523, 1523 (2003).

[6] *See, e.g.*, *id.* at 1525–26 (emphasizing that, dating to the Holocaust, hateful and racist propaganda have been used to demean various ethnic groups); Friedrich Kübler, *How Much Freedom for Racist Speech?: Transnational Aspects of a Conflict of Human Rights*, 27 HOFSTRA L. REV. 335, 336 (1999) (indicating that hate speech regulation in Germany began as a "complex response to the darkest chapter in German history"); Mariana Mello, Note, Hagan v. Australia*: A Sign of the Emerging Notion of Hate Speech in Customary International Law*, 28 LOY. L.A. INT'L & COMP. L. REV. 365, 376 (2006) (citing to Mari J. Matsuda, *Public Response to Racist Speech: Considering the Victim's Story,* 87 MICH. L. REV. 2320 (1989), for the proposition that hate speech can be a "tool for subordination").

[7] Chaplinsky v. N.H., 315 U.S. 568, 572 (1942) (describing fighting words as "those which by their very utterance inflict injury or tend to incite an immediate breach of the peace").

its decisions *Brandenburg v. Ohio*[8] and *Dennis v. U.S.*, whose earlier holding on a similar issue was not completely overruled by *Brandenburg*.[9] Part 2 will conclude by highlighting the Court's continual deference to speech rights, even where the speech is discriminatory, through an examination of *R.A.V. v. City of St. Paul*[10] and successive decisions.

In Part 3, I will examine the ways in which United Nations (UN) instruments, such as the Universal Declaration of Human Rights (UDHR)[11] and the International Covenant on Civil and Political Rights (ICCPR),[12] as well as international hate speech jurisprudence, including *Prosecutor v. Nahimana, Barayagwiza, & Ngeze* or the *Media Case* following the 1994 genocide in Rwanda[13] and other significant decisions in Canada, the United Kingdom, Germany, and Australia, contribute to the current international standard for hate speech regulation. I will demonstrate that, at best, international attitudes towards regulating hate speech are inconsistent and that a precisely articulated standard does not yet exist in the international community.

In Part 4, I will look to aspects of the hate speech regulation standard in the United States in order to recommend a more precise structure for the international standard. To do this, I will address the factors on which the U.S. Supreme Court relied in

---

8 *See Brandenburg*, 395 U.S. 444 (providing the current standard the U.S. Supreme Court uses in evaluating government regulation of advocacy directly related to violence).

9 Dennis v. U.S., 341 U.S. 494 (1951) (distinguishing mere teaching from active advocacy and determining, in that particular instance, that advocating to overthrow the government was a "clear and present danger" to the government and thus could not be protected under the First Amendment).

10 R.A.V. v. City of St. Paul, 505 U.S. 377, 379 (1992) (invalidating a bias-motivated Minnesota statute that sought to criminalize the display of a symbol which "arouses anger, alarm or resentment in others on the basis of race, color, creed, religion or gender" (citing ST. PAUL, MINN. LEGIS. CODE § 292.02 (1990)) due to its overbroad scope and impermissible regulation of speech based on content).

11 Universal Declaration of Human Rights, G.A. Res. 217 (III) A, U.N. Doc. A/RES/217(III) at 71 (Dec. 10, 1948) [hereinafter UDHR].

12 International Covenant on Civil and Political Rights, Dec. 19, 1966, 999 U.N.T.S. 171 [hereinafter ICCPR].

13 Prosecutor v. Nahimana, Case No. ICTR-99-52-T, Judgement and Sentence (Dec. 3, 2003), *available at* http://www.unhcr.org/cgi-bin/texis/vtx/refworld/rwmain?docid=404468bc2 (prosecuting the principal members of several Rwandan media outlets for inciting genocide through the transmission of hate speech).

Brandenburg,[14] as well as the portion of the *Dennis* holding that the *Brandenburg* standard does not directly overrule. I will examine the potential impact of this oversight on international hate speech regulation. I will ultimately recommend a more consistent international framework for hate speech regulation that borrows, in part, from the specificity of the American standard.

2. CURRENT TREATMENT OF HATE SPEECH IN THE UNITED STATES

Freedom of speech under the First Amendment of the Constitution is a sacred right. Grounds for its protection include the (1) preservation of diversity of thought and the individual search for truth in the marketplace of ideas,[15] (2) promotion of self-government,[16] and (3) encouragement of self-expression and individual autonomy.[17] The primary purpose of the First Amendment is to prevent the government from suppressing speech in a way that would inhibit a community's ability to engage in the free exchange of ideas. As Justice Jackson eloquently stated in *West Virginia State Board of Education v. Barnette*, "If there is any fixed star in our constitutional constellation, it is that no official, high or petty, can prescribe what shall be orthodox in politics, nationalism, religion, or other matters of opinion . . . ."[18]

As a general matter, the Court has provided broad protection of individual speech rights, regardless of whether the content of

[14] *See infra* Part 2.1, notes 33–37 and accompanying text.

[15] *See* JOHN STUART MILL, ON LIBERTY AND UTILITARIANISM 60 (1993) ("[I]f any opinion is compelled to silence, that opinion may, for aught we can certainly know, be true. To deny this is to assume our own infallibility.").

[16] *See* Alexander Meiklejohn, *The First Amendment is an Absolute*, 1961 Sup. Ct. Rev. 245, 263 (1961) ("In my view, 'the people need free speech' because they have decided, in adopting, maintaining and interpreting their Constitution, to govern themselves rather than to be governed by others.").

[17] *See* Wooley v. Maynard, 430 U.S. 705, 715 (1977) (invalidating a New Hampshire law mandating the display of the state motto on automobile license plates and in doing so, affirming the First Amendment "right of individuals to hold a point of view different from the majority and to refuse to foster, in the way New Hampshire commands, an idea they find morally objectionable").

[18] W. Va. State Bd. of Educ. v. Barnette, 319 U.S. 624, 642 (1943). In this decision, the Court noted that in order to fully embrace "intellectual individualism and . . . cultural diversities," society must bear the "price of occasional eccentricity and abnormal attitudes." The true "test of [the] substance [of the First Amendment] is the right to differ as to things that touch the heart of the existing order." *Id.* at 641–42.

the speech is unfavorable.[19] Accordingly, the Court has consistently held that government regulation of pure speech is justified *only* if the speech falls within a category that is expressly unprotected by the First Amendment.[20] Precedent dictates that these categories include, primarily: (1) fighting words,[21] (2) true threats,[22] (3) direct incitement,[23] (4) obscenity,[24] (5) child pornography,[25] and (6) deliberate defamation or libel.[26] In other words, hate speech that directly incites another to violence or that qualifies as a true threat—that is, speech that puts an individual in fear of his life or safety[27]—would likely be proscribable under the

---

[19] U.S. v. Schwimmer, 279 U.S. 644, 654-55 (1929) (Holmes, J., dissenting) ("[I]f there is any principle of the Constitution that more imperatively calls for attachment than any other it is the principle of free thought—not free thought for those who agree with us but freedom for the thought that we hate.").

[20] *See* Chaplinsky v. N.H., 315 U.S. 568, 571-72 (1942) ("There are certain well-defined and narrowly limited classes of speech, the prevention and punishment of which have never been thought to raise any Constitutional problem.").

[21] *See id.* at 572 (holding that some forms of expression have little social value and do not communicate ideas and are thus not afforded First Amendment protection).

[22] *See* Virginia v. Black, 538 U.S. 343 (2003) (noting that, although the court struck down a Virginia statute banning cross burning on its face in this instance, a state can ban cross burning *with intent to intimidate* without violating First Amendment speech protections).

[23] *See* Brandenburg v. Ohio, 395 U.S. 444, 447 (1969) (articulating the principal that the "constitutional guarantees of free speech and free press" do not extend to "advocacy [that] is directed to inciting or producing imminent lawless action and is likely to incite or produce such action").

[24] *See, e.g.*, Miller v. Cal., 413 U.S. 15, 24 (1973) (outlining the current standard for identifying obscenity as unprotected expression); Roth v. U.S., 354 U.S. 476, 484 (1957) (holding that obscene expression is "utterly without redeeming social importance").

[25] *See* N.Y. v. Ferber, 458 U.S. 747, 764 (1942) ("When a definable class of material . . . bears so heavily and pervasively on the welfare of children engaged in its production, . . . the balance of competing interests is clearly struck and . . . it is permissible to consider these materials as without the protection of the First Amendment.").

[26] *See* N.Y. Times v. Sullivan, 376 U.S. 254, 279-80 (1964) (identifying "actual malice," or "knowledge that [the defamatory statement] was false or with reckless disregard of whether it was false or not," as the standard by which public officials can recover on defamation or libel claims).

[27] *See* Planned Parenthood of the Columbia/Williamette, Inc. v. American Coalition of Life Activists, 290 F.3d 1058, 1076-77, 1088 (9th Cir. 2002) (defining a "true threat" as a "statement which, in the entire context and under all the circumstances, a reasonable person would foresee would be interpreted by those to whom the statement is communicated as *a serious of intent to inflict bodily harm* up-

First Amendment. On the other hand, the government would *not* be justified in banning hate speech due merely to its hateful content.[28] The Court fears that such content-based bans will encourage future proscription of constitutionally legitimate speech.[29]

2.1. *Incitement and Hate Speech Under* Brandenburg *and* Dennis

Advocates of restricting hate speech argue that the current U.S. standard is too speech-protective and that hate speech by its content alone spurs more hate, which is likely to lead to mass discrimination and violence.[30] In *Gitlow v. New York*, Justice Holmes countered that notion by stating that "[e]very idea is an incitement"[31]—that, without a more speech-protective standard, all speech, especially speech that expresses views outside the mainstream, may be curtailed based merely on its potential to cause danger.[32] The Court's approach to regulating speech that incites violence offers insight into when it is appropriate to restrict hate speech on grounds that hate speech will trigger violent action.

In *Brandenburg*, the Court outlined the standard by which to assess the constitutionality of a government restriction on speech

---

on that person" (emphasis added), and holding that intimidating physicians from providing reproductive health services was an unconstitutional exercise of the American Coalition of Life Activists' First Amendment rights).

[28] *See* R.A.V. v. City of St. Paul, 505 U.S. 377, 392–93 (1992) (holding it impermissible to ban hate speech based merely on its message and stating that the state had "no such authority to license one side of a debate to fight freestyle, while requiring the other to follow Marquis of Queensberry rules").

[29] *See* Toni M. Massaro, *Equality and Freedom of Expression: The Hate Speech Dilemma*, 32 WM. & MARY L. REV. 211, 215 (1991) (emphasizing the difficulty of determining the precise parameters of racist and homophobic speech and the consequent likelihood that content-based restrictions of hate speech could be overly broad).

[30] *See* Thomas J. Webb, *Verbal Poison—Criminalizing Hate Speech: A Comparative Analysis and a Proposal for the American System*, 50 WASHBURN L.J. 445, 445-46 (2011) (outlining the potentially destructive effects of hate speech, including "distress, intimidation, and fear," and noting that "the United States, in effect, has become a safe haven for the promotion of hate speech").

[31] Gitlow v. New York, 268 U.S. 652, 673 (1925).

[32] *See* Nadine Strossen, *Incitement to Hatred: Should There Be a Limit?*, 25 S. ILL. U. L. J. 243, 250 (2001) (detailing historical arguments for restricting speech only in cases of imminent and substantial danger and explaining how these arguments shape the current speech-protective standard for hate speech regulation in the United States).

that incites violence. In that decision, the Court determined that speech advocating violence may be proscribed *only if* the speaker intends to incite or produce actions that are imminent and lawless.[33] In addition, the speech must be objectively likely to produce a violation.[34] Accordingly, the Court has distinguished abstract teachings or generalized advocacy to engage in violent action from speech that is intended to produce an imminent violation and is likely to do so.[35] Consistent with Justice Black's dissent in *Dennis*,[36] *Brandenburg* stands for the proposition that the mere advocacy of illegal action is not enough to justify the suppression of speech.[37]

Even in cases such as *NAACP v. Claiborne Hardware*, where the speech is especially inflammatory,[38] the Court has erred on the side of protecting individual speech rights.[39] The *Claiborne Hardware* decision clarifies the Court's stance on acts of violence within a mass protest: when a protest, set in a speech-rich context, is intermingled with isolated incidents of violence or penalty-inducing behavior, only those who commit the acts of violence are

---

[33] *See* Brandenburg v. Ohio, 395 U.S. 444, 447-49 (1969) ("[T]he constitutional guarantees of free speech and free press do not permit a State to forbid or proscribe advocacy of the use of force or of law violation except where such advocacy is directed to inciting or producing imminent lawless action and is likely to incite or produce such action.").

[34] *See id.* at 448 (requiring that states distinguish between "the mere abstract teaching . . . of the moral propriety or even moral necessity for a resort to force and violence" and "preparing a group for violent action and steeling it to such action" (quoting Noto v. U.S., 367 U.S. 290, 297-98 (1961))).

[35] *Id.*

[36] *See* Dennis v. U.S., 341 U.S. 494, 580 (1951) (Black, J., dissenting) ("[T]he chief reason for jettisoning the rule is the expressed fear that advocacy of Communist doctrine endangers the safety of the Republic. Undoubtedly, a governmental policy of unfettered communication of ideas does entail dangers. To the Founders of this Nation, however, the *benefits derived from free expression were worth the risk.*") (emphasis added).

[37] *See* Brandenburg, 395 U.S. at 448–49 (holding that a "statute which fails to draw this distinction [between mere advocacy of illegal action and advocacy intended to incite imminent lawlessness] impermissibly intrudes upon the freedoms guaranteed by the First and Fourteenth Amendments").

[38] *See* NAACP v. Claiborne Hardware Co., 458 U.S. 886, 902 (1982) (quoting speaker Charles Evers, with regard to a boycott of white merchants in Mississippi: "If we catch any of you going in any of them racist stores, we're gonna [sic] break your damn neck.").

[39] *See id.* at 928 ("An advocate must be free to stimulate his audience with spontaneous and emotional appeals for unity and action in a common cause.").

liable for their actions.[40] The Court in *Claiborne Hardware* determined that the speaker Charles Evers' speech, regardless of its tone and message, did *not* qualify as proscribable incitement under *Brandenburg*; instead, it was merely coercive advocacy.[41]

2.2. *A Speech-Protective Framework for Hate Speech Regulation in* R.A.V. v. City of St. Paul *and* Virginia v. Black

The majority in *Claiborne Hardware* stated that when assessing the validity of state regulation of individual speech rights under the First Amendment, the Court should look first for "precision of regulation."[42] Accordingly, the Court typically assesses hate speech restrictions with the utmost care, due to the possibility that they may stymie legitimate speech in the future. The *R.A.V.* and *Virginia v. Black* decisions illustrate this concern. Decided in 1992, *R.A.V.* reflects the Court's standing methodology when analyzing hate speech restrictions. The St. Paul Bias-Motivated Crime Ordinance at issue in *R.A.V.* provided:

> Whoever places on public or private property a symbol, object, appellation, characterization or graffiti, including, but not limited to, a burning cross or Nazi swastika, which one knows or has reasonable grounds to know arouses anger, alarm or resentment in others on the basis of race, color, creed, religion or gender commits disorderly conduct and shall be guilty of a misdemeanor.[43]

The Court, while accepting the Minnesota Supreme Court's limiting construction of the statute to reach only fighting words,

---

[40] *See id.* at 916-18 (emphasizing that the state may impose liability *only* on those who engage in violent conduct or who are directly responsible for the consequences of such conduct and *not* on those who engage in protected speech activity).

[41] *See id.* at 927–28 (indicating that Evers' speech constituted neither fighting words under *Chaplinsky* nor incitement under *Brandenburg*, and concluding that "[t]he emotionally charged rhetoric of . . . Evers' [speech] did not transcend the bounds of protected speech").

[42] *Id.* at 916 (citing NAACP v. Button, 371 U.S. 415, 438 (1963)). *Button* further stands for the proposition that "[b]road prophylactic rules in the area of free expression are suspect" and that "precision of regulation" is most important in "area[s] so closely touching [an individual's] most precious freedoms." *Button*, 371 U.S. at 438.

[43] R.A.V. v. City of St. Paul, 505 U.S. 377, 380 (1992) (citing ST. PAUL, MINN. LEGIS. CODE § 292.02 (1990)).

held the ordinance facially unconstitutional because it sought to restrict a particular brand of fighting words on the basis of content, for reasons *other than* those underlying the decision to strip fighting words of their constitutional protection in the first place.[44] Writing for the majority, Justice Scalia further clarified the distinction by supporting the government's right to "prohibit only that obscenity which is the most patently offensive *in its prurience* [or] which involves the most lascivious displays of sexual activity,"[45] but refusing to endorse viewpoint-based discrimination[46] in speech regulations.

Justice Sandra Day O'Connor's majority in *Black* followed Justice Scalia's reasoning in *R.A.V.*[47] when it held that Virginia could proscribe cross burning if the actor intended to intimidate because cross burning is a particularly vicious form of intimidation.[48] It appears then that under the *R.A.V.-Black* scheme, the government may proscribe hate speech based on its content only when it (1) falls into a previously unprotected category under

---

[44] *See id.* at 391 (concluding that, since "the ordinance applies only to 'fighting words' that insult, or provoke violence 'on the basis of race, color, creed, religion or gender,'" the regulation is "facially unconstitutional" because it seeks to eliminate certain unfavorable viewpoints, as opposed to certain particularly vicious kinds of fighting words, the regulation is impermissible).

[45] *Id.* at 388.

[46] In other words, the government *can* impose restrictions of certain fighting words for being particularly likely to incite violence or certain brands of obscenity for being particularly lewd. It *cannot*, however, impose restrictions based on viewpoint, e.g. a ban on only racist fighting words, regardless of how disfavored the viewpoint is. *See* Brown v. Entm't Merchants Ass'n, 131 S. Ct. 2729, 2732 (2011) (invalidating a California statute banning the sale of violent video games on the grounds that it was "wildly underinclusive, raising serious doubts about whether the State is pursuing the interest it invokes or is instead disfavoring a particular speaker or viewpoint"). *See generally* Snyder v. Phelps, 131 S. Ct. 1207, 1217 (2011) (articulating that even highly offensive speech concerning general or public issues—here, signs containing messages such as "You're Going to Hell" and "God Hates You"—deserves full First Amendment protection).

[47] *See* Virginia v. Black, 538 U.S. 343, 362 (2003) ("Virginia's statute does not run afoul of the First Amendment insofar as it bans cross burning with intent to intimidate. Unlike the statute at issue in *R.A.V.*, the Virginia statute does not single out for opprobrium only that speech directed toward 'one of the specified disfavored topics.' It does not matter whether an individual burns a cross with intent to intimidate because of the victim's race, gender, or religion, or because of the victim's 'political affiliation, union membership, or homosexuality.'") (quoting *R.A.V.*, 505 U.S. at 391).

[48] *See Black*, 538 U.S. at 363 (highlighting cross burning as a long-standing symbol of imminent violence).

the First Amendment, and (2) represents a particularly virulent strain of the kind of speech in that unprotected category. The Court's two-step analysis constitutes a very precise basis for hate speech regulation. And as the Court has not deemed many categories of speech unprotected under the First Amendment,[49] this foundation is also very speech-protective. Ultimately, however, if the international community's goal for hate speech regulation is to value both the speaker's right to express a viewpoint *and* the listener's right to remain free from discrimination, then adopting the *R.A.V.-Black* framework in its entirety is an incomplete solution.[50]

As the following Part explains, the international community gives considerable weight to the non-violent yet harmful import of hate speech. Though this priority is a noble backdrop for regulation, it holds perhaps too much potential for overregulation. Thus, the international standard for hate speech regulation may benefit from internalizing the precision, if not the near-absolute protection of individual speech rights, of the American framework.[51]

3. CURRENT TREATMENT OF HATE SPEECH IN INTERNATIONAL LAW

International human rights instruments codify the rights to both freedom of expression and freedom from discrimination. Accordingly, the relationship between the rights of the speaker and the listener plays a larger role in protecting individual speech under the international approach than it does under the U.S. approach.[52] The ICCPR, among many conventions and treaties,

---

[49] *See supra* notes 21-27 and accompanying texts.

[50] *See* Webb, *supra* note 30, at 446 (noting that, under the current American standard for hate speech regulation, "hate speech cannot be criminalized when it is simply encouraging hatred, which can be as harmful as expressly inciting violence or threatening others").

[51] *See* Massaro, *supra* note 29, at 243 (outlining drawbacks to the international approach to hate speech regulation by highlighting the "contested interpretation of equality" and noting that the role of government is not to "take a side in intergroup hostilities, [albeit] if only to even the score," but "to remain neutral when policing intergroup conflicts").

[52] *See, e.g.*, UDHR, *supra* note 11, at art. 29(2) ("In the exercise of his rights and freedoms, everyone shall be subject only to such limitations as are determined by law solely for the purpose of securing due recognition and respect for the rights and freedoms of others and of meeting the just requirements of morality, public order and the general welfare in a democratic society."); Scott J. Catlin, *A Proposal*

addresses this balance.[53] *Prosecutor v. Nahimana* and other landmark decisions demonstrate, however, that in a herculean effort to strengthen two often competing fundamental human rights, the international community is no closer to articulating a consistent standard for prosecuting hate speech.[54]

3.1. An Interplay of Rights: Freedom of Expression and Freedom from Discrimination as Codified in International Human Rights Instruments

Many international legal instruments explicitly designate the right to free expression as a fundamental human right that is worthy of protection. Article 19 of both the UDHR and the ICCPR highlights the "right to hold opinions without interference"[55] as well as the right to receive a wide breadth of information through any media.[56] Article 13 of the Convention on the Rights of the Child (CRC) further protects a child's freedom to "seek, receive and impart information and ideas of all kinds."[57] In addition, Article 4 of the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) requires state parties to

---

for Regulating Hate Speech in the United States: Balancing Rights Under the International Covenant on Civil and Political Rights, 69 NOTRE DAME L. REV. 771, 795 (1993–1994) (emphasizing that the fundamental tenets of international human rights agreements "require a balancing of the speaker's right to free speech against the listener's right to have her inherent human dignity protected from hate speech injuries"); Elizabeth F. Defeis, *Freedom of Speech and International Norms: A Response to Hate Speech*, 29 STAN. J. INT'L L. 57, 71 (1992–1993) ("The rights of equality and non-discrimination are central in the Universal Declaration and no rights, including speech rights, may be asserted to destroy them.").

[53] *See* ICCPR, *supra* note 12, at arts. 19(2) ("Everyone shall have the right to freedom of expression") and 20(2) ("Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.").

[54] *See generally* Alexander Zahar, *The ICTR's "Media" Judgment and the Reinvention of Direct and Public Incitement to Commit Genocide*, 16 CRIM. L.F. 33, 47–48 (2005), *available at* http://ssrn.com/abstract=1348507 (arguing that the Media judgment failed to articulate solid reasons as to why the broadcasted and written statements are at issue, despite publicly preaching hate, constituted direct and public incitement to genocide).

[55] ICCPR, *supra* note 12, at art. 19(1); UDHR, *supra* note 11, at art. 19.

[56] ICCPR, supra note 12, at art. 19(2); UDHR, *supra* note 11, at art. 19.

[57] Convention on the Rights of the Child art. 13, Nov. 20, 1989, 1577 U.N.T.S. 3.

criminalize "all dissemination of ideas based on racial superiority or hatred, [and] incitement to racial discrimination."[58]

These instruments also support the notion that the right to free expression does not stand alone. Rather, it is measured in relation to other fundamental human rights, including the right to freedom from discrimination. Article 19(3) of the ICCPR indicates that freedom of expression is contingent on the fulfillment of "special duties and responsibilities" and may be restricted on grounds of, most notably, "respect of the *rights . . . of others*."[59] This qualification is a reflection of the preambles to both the UDHR and ICCPR, which emphasize the importance of recognizing and maintaining basic dignity and equality among all individuals.[60] In addition, Article 26 of the ICCPR states that "all persons are equal before the law and are entitled *without any discrimination* to the equal protection of the law."[61]

Article 20(2) of the ICCPR addresses hate speech directly, stating that all "advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law."[62] Although it invokes the incitement language of *Brandenburg* and *Dennis*, Article 20 broadens the range of proscribable speech based on content to include not only speech that incites violence, but also speech that incites discrimination and hostility.[63] Further, the ICCPR obligates state parties to adopt laws against this kind of speech. This provision is outlined in Article 2:

---

[58] International Convention on the Elimination of All Forms of Racial Discrimination art. 4(a), Mar. 7, 1966, 5 I.L.M. 352, 660 U.N.T.S. 195.

[59] ICCPR, *supra* note 12, at art. 19(3) (emphasis added).

[60] *See id.* pmbl. ("[r]ealizing that the individual [has] duties to other individuals and to the community to which he belongs . . . ."); UDHR, *supra* note 11, pmbl. ("[T]he peoples of the United Nations have in the Charter reaffirmed their faith in . . . the dignity and worth of the human person and in the equal rights of men and women . . . .").

[61] ICCPR, *supra* note 12, at art. 26 (emphasis added).

[62] *Id.* at art. 20(2).

[63] *See* Catlin, *supra* note 52, at 799 (detailing Article 20's broad reach in its prohibition of incitement to discrimination and hostility, and indicating that this feature of the international approach is most at odds with the American approach to regulating hate speech); *see generally* David Filvaroff et al., *The Substantive Rights and United States Law*, *in* U.S. RATIFICATION OF THE INTERNATIONAL COVENANTS ON HUMAN RIGHTS 54 (Hurst Hannum & Dana D. Fischer eds., 1993).

> Where not already provided for by existing legislative or other measures, each State Party to the present Covenant undertakes to take the necessary steps, in accordance with its constitutional processes and with the provisions of the present Covenant, to adopt such legislative or other measures as may be necessary to give effect to the rights recognized in the present Covenant.[64]

The ICCPR's provision regarding incitement to hostility and discrimination are not sufficiently speech-protective to square with the American standard.[65] As a result, the United States made reservations to its ratification of the ICCPR. In order to maintain ultimate discretion with regard to speech regulation, the U.S. Senate declared the ICCPR a non-self-executing treaty, or one that American courts cannot directly enforce until Congress and the Executive Branch pass appropriate legislation.[66]

3.2. Regulating Hate Speech in the Context of Incitement to Genocide in Rwanda: A Reminder of the Dangers of Unchecked Expression

Unlike in other contexts in international law, hate speech regulation in the context of genocide is similar to the American system of regulating incitement. In its 2003 *Prosecutor v. Nahimana* decision following the Rwandan genocide, the International Criminal Tribunal of Rwanda (ICTR) adopted a precise standard for prosecuting hate speech that incorporated many of the speech-protective elements that compose the current American standard for regulation of incitement.[67] In that case, the ICTR was right to be speech protective; it exercised its power to convict an individual for a genocide-related crime based solely on the transmission of hateful speech with appropriate caution. Uniquely, the ICTR was

---

[64] ICCPR, *supra* note 12, at art. 2.

[65] *See* Catlin, *supra* note 52, at 799 (noting that the international approach, due to its broader parameters, necessitates more content-based regulations, many of which might be impermissible under the American standard for hate speech regulation).

[66] *See id.* at 802 (emphasizing the "dichotomous approach to the implementation of [the United States'] international human rights agreements" and stressing that although the United States maintains a commitment internationally to the tenets of such agreements, it "reserves the option to only partially implement them domestically").

[67] *See infra* notes 82–90 and accompanying texts.

also able to underscore the value that international law places on the right to non-discrimination by ultimately convicting three principal figures in the Rwandan media for incitement to genocide.[68] This result also fits; restricting speech that incites genocide—that is, prioritizing the right to freedom from discrimination when that discrimination takes the form of ethnic cleansing—is legitimate, if only to deter future occurrences.

However, *Nahimana* and related post-genocide decisions stand separately. Genocide is a horror of such magnitude that it validates arguments on both sides of speech regulation.[69] In less dire cases, it is more difficult to equally prioritize the rights to freedom of expression and freedom from discrimination. Courts around the world have thus shown varying degrees of prudence when proscribing hate speech. Nonetheless, it is helpful to examine the ICTR's decisions for their analytical clarity and illumination of the power of hate speech and the worst consequences of allowing absolute freedom of expression.

3.2.1. Background of the Rwandan Genocide

Current President of the International Association of Genocide Scholars, Professor William A. Schabas, wrote: the "road to genocide in Rwanda was paved with hate speech."[70] Beginning in the 1930s and spurred by Belgian colonialism, the ethnic divide between the Hutu and Tutsi races in Rwanda hardened over six decades leading to the 1994 genocide whose 100-day span oversaw the massacre of over one million Tutsis and moderate Hutus.[71]

In 1993, Hutu extremists established a radio station called *Radio-Télévision Libre des Milles Collines* (RTLM),[72] which was

---

[68] *See infra* notes 78–81 and accompanying texts.

[69] *See* Catharine A. MacKinnon, *Prosecutor v. Nahimana, Barayagwiza, & Ngeze*, 98 AM. J. INT'L L. 325, 330 (2004) ("The ICTR's pathbreaking ruling shows that equality and speech rights can be harmonized when courts face the power of media to kill.").

[70] William A. Schabas, *Hate Speech in Rwanda: The Road to Genocide,* 46 MCGILL L.J. 141, 144 (2000).

[71] *See* Scott Straus, *How Many Perpetrators Were There in the Rwandan Genocide? An Estimate,* 6 J. GENOCIDE RES. 85 (2004).

[72] *See* Alison Des Forges, *Call to Genocide: Radio in Rwanda, 1994*, *in* THE MEDIA AND THE RWANDA GENOCIDE 41, 44 (Allan Thompson ed., 2007) (detailing the harmless beginnings of RTLM and emphasizing that the "station was meant to be the voice of the people . . . ").

supported by the government-controlled Radio Rwanda, to broadcast messages of hate. Prior to the creation of RTLM and the start of the Tutsi extermination, the French-language tabloid *Kangura*,[73] whose primary demographic was the illiterate Hutu population, advocated the killing of Tutsis.[74] Fear-mongering through threats and rumors played a significant role in inciting hatred of Tutsis in the local population.[75]

The prevalence of hate media intensified following the death of Rwandan President Habyarimana in 1994. RTLM broadcasts incited Hutus to eliminate the *inyenzi* or "cockroach,"[76] which became a euphemism for the Tutsi people during the genocide.[77] Despite the damaging effects of hate media in Rwanda, the United States declined intervention, reasoning in part that shutting down the broadcast of these hateful messages would impermissibly interfere with the sovereignty of the Rwandan government and would bear an uncomfortable resemblance to censorship.[78]

---

[73] *See* Zahar, *supra* note 54, at 45 (characterizing *Kangura* as "a Hutu-nationalist rag, hateful of the [Tutsi-created Rwandan Patriotic Front or] RPF and the threat it posed to what Ngeze saw as the glorious Hutu revolution of 1959, when the masses threw off the Tutsi yoke, founded the Republic, regained their dignity, and set Rwanda on the path to modernity").

[74] *Id.* (indicating that *Kangura* aimed its advocacy towards illiterate Rwandans).

[75] *See* Des Forges, *supra* note 72, at 45 (documenting the RTLM's turn to sensationalism in 1993 and its twin aims of "underlin[ing] supposed Tutsi brutality and heighten[ing] Hutu fears of Tutsi").

[76] *See* Prosecutor v. Nahimana, Case No. ICTR-99-52-T, Judgement and Sentence, iv (Dec. 3, 2003), *available at* http://www.unhcr.org/cgi-bin/texis/vtx/refworld/rwmain?docid=404468bc2, (defining *inyenzi* as: "[c]ockroach; group of refugees set up in 1959 to overthrow the new regime; sympathizer of RPF; sometimes used to refer to Tutsi").

[77] *See* Prosecutor v. Ruggiu, Case No. ICTR-97-32-I, Judgement and Sentence, ¶ 44 (June 1, 2000), *available at* http://www.unictr.org/Portals/0/Case%5CEnglish%5CRuggiu%5Cjudgement%5Crug010600.pdf (convicting Georges Ruggiu, a Belgian broadcaster who pled guilty to committing incitement to genocide in Rwanda and who admitted that by 1994 "the term 'Inyenzi' [had become] synonymous with the term 'Tutsi' . . . [and had come] to designate the Tutsis as 'persons to be killed'").

[78] *See* Gregory H. Stanton, *The Rwandan Genocide: Why Early Warning Failed*, 1 J. AFR. CONFLICTS AND PEACE STUD. 6, 9 (2012), *available at* http:// scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1015&context=jacaps (indicating that the U.S. Ambassador to Rwanda, David Rawson, and the State Department defended RTLM's right to broadcast as freedom of speech) (internal quotations omitted).

3.2.2. Prosecuting Hate in the Aftermath of the Genocide

The *Nahimana* decision marked "international criminal law's first reexamination of the link between mass media and mass slaughter" since the Nuremburg Trials.[79] In that case, the ICTR addressed the accountability of Rwandan hate media outlets in the genocide. The Trial Chamber found three Rwandan journalists of radio and print media, Ferdinand Nahimana, Jean-Bosco Barayagwiza, and Hassan Ngeze, guilty of several counts under the 1948 Convention for the Prevention and Punishment of the Crime of Genocide ("Genocide Convention"), including direct and public incitement to commit genocide.[80] These convictions on the count of incitement to commit genocide[81] indicate that hate speech can be sufficient to constitute one of international law's most atrocious crimes.[82]

The ICTR previously addressed the crime of direct and public incitement to commit genocide in its 1998 decision, *Prosecutor v. Jean-Paul Akayesu*.[83] This decision considers the implications of the terms 'public' and 'direct,' as well as the *mens rea* required to incur

---

[79] Recent Case, *International Law — Genocide — U.N. Tribunal Finds That Mass Media Hate Speech Constitutes Genocide, Incitement to Genocide, and Crimes Against Humanity. — Prosecutor v. Nahimana, Barayagwiza, and Ngeze (Media Case), Case No. ICTR-99-52-T (Int'l Crim. Trib. for Rwanda Trial Chamber I Dec. 3, 2003)*, 117 Harv. L. Rev. 2769 (2004) [hereinafter *Recent Case: The Media Case*].

[80] *See* Prosecutor v. Nahimana, Case No. ICTR-99-52-T, Judgement and Sentence, ¶¶ 1091-1094 (Dec. 3, 2003), *available at* http://www.unhcr.org/cgi-bin/texis/vtx/refworld/rwmain?docid=404468bc2.

[81] Note that in its 2007 decision, the Appeals Chamber affirmed the convictions of Nahimana and Ngeze for direct and public incitement to commit genocide. *See* Prosecutor v. Nahimana, Case No. ICTR-99-52-A, Appeal Judgement (Nov. 28, 2007), *available at* http://www.haguejusticeportal.net/Docs/Court%20Documents/ICTR/Nahimana%20et%20al_Appeal%20Judgment.pdf (supporting the Trial Chamber's distinction between general hate speech and incitement to commit genocide, *id.* ¶ 715, and upholding the Trial Chamber's convictions of Nahimana and Ngeze for direct and public incitement to commit genocide, *id.* ¶¶ 1051, 1113).

[82] *See* MacKinnon, *supra* note 69, at 328-29 ("*The Media Case* is notable for holding a newspaper editor and a broadcast executive criminally accountable not only for the crime of what they said, but for the crimes their words *did* . . . .").

[83] Prosecutor v. Akayesu, Case No. ICTR-96-4-T, Judgement (Sept. 2, 1998), *available at* http://www.unictr.org/Portals/0/Case%5CEnglish%5CAkayesu%5Cjudgement%5Cakay001.pdf.

criminal liability.[84] 'Public' refers to a "call for criminal action to a number of individuals" by means of a speech, public threats, radio and television broadcasts, or communication through other mass media.[85] 'Direct' refers to incitement that is specifically intended to elicit a response, in contrast with "vague or indirect suggestion."[86] In order for the incitement to be direct, the target audience must also be objectively likely to act on it. In other words, the audience must understand the speech at issue, and the speech must be persuasive.[87] Both *Akayesu* and the *Nahimana* appeal indicate that the presence of genocidal intent[88] alone is sufficient to constitute a crime, regardless of whether the speech at issue successfully triggered genocide.[89]

The *Akayesu* analysis resembles the standard for proscribing hate speech that constitutes incitement under current U.S. law.[90] Similarly, the ICTR's evaluation in *Nahimana* aligns with that of the U.S. Supreme Court in *Brandenburg* and *Dennis*. Through an examination of the specificity and tone of the hate speech at issue and the context of its dissemination, the Tribunal sought to determine whether the speech actively advocated violence or

---

[84] *See id.* ¶¶ 556-60 (outlining the factors considered in identifying incitement to be public or direct and stating that the required mens rea lies in the specific intent "to directly prompt or provoke another to commit genocide").

[85] *Id.* ¶ 556 (citing the International Law Commission's definition of public incitement).

[86] *Id.* ¶ 557.

[87] *See id.* ¶¶ 557-58 ("The Chamber will . . . consider on a case-by-case basis whether, in light of the culture of Rwanda . . . acts of incitement can be viewed as direct or not, by focusing mainly on the issue of whether the persons for whom the message was intended immediately grasped the implication thereof.").

[88] *See* Prosecutor v. Nahimana, Case No. ICTR-99-52-A, Appeal Judgement ¶ 523 (Nov. 28, 2007), *available at* http://www.haguejusticeportal.net/Docs/Court%20Documents/ICTR/Nahimana%20et%20al_Appeal%20Judgment.pdf (citing the definition of genocidal intent in Article 2(2) of the Statute of the International Criminal Tribunal for Rwanda as the "intent to destroy, in whole or in part, a national, ethnical, racial or religious group") (internal quotations omitted).

[89] *See id.* ¶ 678 (emphasizing that "the drafters of the [Genocide] Convention intended to punish direct and public incitement to commit genocide, even if no act of genocide was committed, the aim being to forestall the occurrence of such acts"); Prosecutor v. Akayesu, Case No. ICTR-96-4-T, Judgement, ¶ 561 (Sept. 2, 1998), *available at* http://www.unictr.org/Portals/0/Case%5CEnglish%5CAkayesu%5Cjudgement%5Cakay001.pdf.

[90] *See supra* Part 2.1, notes 33-37 and accompanying text.

simply educated in the abstract.[91] Although largely cabined to circumstances as extreme as genocide, the reasoning in the *Nahimana* decision is an influential example of precision in hate speech regulation.[92]

3.3. *An Inconsistent Treatment of Hate Speech in Recent International Jurisprudence*

As demonstrated by hate speech decisions and regulations in Canada, the United Kingdom, Germany, and Australia, the international standard for hate speech regulation becomes less consistent in the absence of an equalizing circumstance. Depending on the country and its history and culture, the standard vacillates between more and less speech-protection, closer to and further from the American system.

3.3.1. *Canada*

Canada derives its commitment to combating hate speech in part from Germany's history and the dissemination of hateful propaganda during the Holocaust.[93] For example, in marked opposition to the American system of speech regulation, the Canadian Supreme Court in *R. v. Keegstra*[94] sustained the conviction of a teacher charged with making anti-Semitic comments to his students.[95] In that case, the teacher was convicted under a statute that criminalizes the "promotion of hatred . . . towards any section of the public distinguished by color, race, religion or ethnic origin."[96] The statute does not mention

[91] *See Recent Case: The Media Case*, *supra* note 79, at 2772 (emphasizing the Tribunal's consideration of context when determining whether the speech at issue "was intended to promote an offense or merely to educate persuasively").

[92] *See* MacKinnon, *supra* note 69, at 330 ("The strong but subtle principles articulated in *The Media Case*, applicable to many legal areas of speech regulation, will have an impact around the world, not least in countries that have tended to see themselves as exempt from horrors of Rwanda's gravity.").

[93] *See* Kübler, *supra* note 6, at 337 ("[T]he German example has inspired legislation in other countries. The Canadian Rules, for example, have been . . . based on the premise that the successes of modern advertising, the triumph of impudent propaganda such as Hitler's, have qualified sharply [the] belief in the rationality of man.") (internal quotations omitted).

[94] R. v. Keegstra, [1990)] 3 S.C.R. 697 (Can.).

[95] *See id.* at 698.

[96] Criminal Code, R.S.C. 1985, c. C-46, § 319(2) (Can.).

incitement to violence and is similar to the statute in *R.A.V.*, which the U.S. Supreme Court deemed overbroad.[97] However, the Canadian Supreme Court concluded that although the statute appears to violate the principles of Section 2(b) of the Canadian Charter of Rights and Freedoms,[98] it is constitutional under Section 1 of the Charter.[99]

The Court's reasoning in *Keegstra* is based on principles of "individual self-fulfillment and human flourishing."[100] This language resembles First Amendment ideals of the search for truth, tolerance of diversity, and the preservation of robust public debate.[101] However, the Canadian view is that these ideals are best achieved by prioritizing individual dignity and social harmony.[102] The dissemination of hate speech undermines these priorities.

3.3.2. United Kingdom

The United Kingdom does not have a consistent framework for hate speech regulation. Historically, sustaining a conviction on a count of seditious libel, for example, required a showing of intent.[103] The Race Relations Act of 1965,[104] although broad in its definition of incitement, included a similar requirement, which

---

[97] *See supra* notes 10, 43-46 and accompanying texts.

[98] Canadian Charter of Rights and Freedoms, Part I of the Constitution Act, 1982, *being* Schedule B to the Canada Act, 1982, c. 11 (U.K.), at Sec. 2(b) (protecting "freedom of thought, belief, opinion and expression, including freedom of the press and other media of communication").

[99] *Id.* § 1 ("The Canadian Charter of Rights and Freedoms guarantees the rights and freedoms set out in it subject *only to such reasonable limits* prescribed by law as can be demonstrably justified in a free and democratic society.") (emphasis added).

[100] *See Keegstra*, 3 S.C.R. at 728 (citing the Canadian Charter).

[101] *See supra* notes 15-18 and accompanying texts.

[102] *See* Rosenfeld, *supra* note 5, at 1544 (underscoring that under the Canadian hate speech regulation standard, the transmission of hate speech is "more dangerous than its suppression as it is seen as likely to produce enduring injuries to self-worth and to undermine social cohesion in the long run").

[103] *See* Stanley Halpin, *Racial Hate Speech: A Comparative Analysis of the Impact of International Human Rights Law upon the Law of the United Kingdom and the United States*, 94 MARQ. L. REV. 463, 467-68 (2010-2011) (detailing past seditious libel laws in the United Kingdom that required "the element of intentional promotion of actual violence to be present").

[104] RACE RELATIONS ACT, 1965, c. 73, § 6 (1) (Eng.).

made it difficult to prosecute hate speech.[105] Over the years, a number of statutory provisions have emerged, broadening the United Kingdom's commitment to restricting hate speech. In 1986, hate speech amounting to harassment of a group became punishable under the Section 5 of the Public Order Act.[106] The enactment of the Protection from Harassment Act followed in 1997.[107] In 2006, the Racial and Religious Hatred Act made it a criminal offense in England and Wales to publish, broadcast, or otherwise disseminate hate speech targeting religious groups.[108]

The overall effectiveness of these statutes, however, remains unclear.[109] As a general matter, the judicial system does not appear to favor hate speech prosecutions, particularly when that speech is unlikely to incite actual violence.[110] In addition, although there is a relatively high incidence of reporting hate speech crimes, a large number of prosecutions are dropped.[111]

3.3.3. Germany

Germany maintains a particularly strong commitment to regulating hate speech due to the virulent strain of hate propaganda perpetuated by the Nazis that led, ultimately, to the Holocaust. As in Canada, German courts view the right to self-expression as part of a broader bundle of individual rights and

---

[105] *See* Rosenfeld, *supra* note 5, at 1546 (noting that although the Race Relations Act of 1965 centered on "incitement to *hatred* rather than . . . incitement to violence," it did have an intent requirement) (emphasis added). Note that the Race Relations Act of 1965 was amended by the Race Relations Act of 1976, and the intent requirement was dropped.

[106] PUBLIC ORDER ACT, 1986, c. 64, §§ 5-6 (Eng.).

[107] PROTECTION FROM HARASSMENT ACT, 1997, c. 40, § 7 (Eng.).

[108] RACIAL AND RELIGIOUS HATRED ACT, 2006 (Eng.).

[109] *See* Rosenfeld, *supra* note 5, at 1547 (arguing that the statutes discussed "provide more tools in the British legal arsenal against hate speech, but have not thus far led to any clearer or more definitive indication of the ultimate boundaries of punishable hate speech in the United Kingdom").

[110] *See* Halpin, *supra* note 103, at 469 (tracing the declining number of hate speech prosecutions in the United Kingdom).

[111] *See id.* at 473 (citing MGMT. INFO. BRANCH, CROWN PROSECUTION SERV., RACIST AND RELIGIOUS INCIDENT MONITORING (2007) to illustrate the high percentages of dropped prosecutions and the "few violations of the 1998 Hate Crime Act [that are] identified as accepted for prosecution").

duties to the community.[112] Today, the German government may impose criminal liability for speech that inflicts harm on human dignity or targets particular minority groups, based on their religious or ethnic origins.[113] Furthermore, in contrast to the American standard for hate speech regulation, the requirement that hate speech must meet an incitement threshold is absent or very minimal under the German standard.[114]

3.3.4. Australia

In Australia, unlike in other countries, the right to free expression is a strictly common law notion.[115] As a result, every state has its own set of restrictions on discriminatory or derogatory speech.[116] These state regulations and federal law[117] appear to specifically target victimization based on race, and the most common avenues of redress are civil penalties. The Australian government rarely enforces the few criminal hate speech regulations that currently exist.[118] One reason for this may be that the focus on civil remedies represents a more comprehensive system of regulation, as criminal statutes are typically more narrowly construed.[119]

Even in the absence of a codified right to free speech, however, Australian courts have sought to uphold speech rights. To

[112] *See* Rosenfeld, *supra* note 5, at 1549-50 (delineating Germany's commitment to preserving the balance "between the self-expression needs of speakers and the self-respect and dignity of listeners").

[113] *See id.* at 1551 (explaining that the German government has a wide variety of legal tools it uses to combat hate speech).

[114] *See id.* at 1551-52 (outlining the lower standards of proof that are required in Germany, with specific reference to prohibitions on denials regarding the Holocaust).

[115] Katharine Gelber, *Hate Speech and the Australian Legal and Political Landscape*, *in* HATE SPEECH AND FREEDOM OF SPEECH IN AUSTRALIA 2, 4-5 (Katharine Gelber & Adrienne Stone eds., 2007) (contrasting the implied right to free speech in Australia with the explicit right to free speech in other countries, such as the United States and the United Kingdom).

[116] *See id.* at 5 (listing Australia's various anti-discrimination laws by jurisdiction).

[117] *See, e.g.*, *Racial Discrimination Act 1975* (Cth) s 18C (Austl.) (prohibiting offensive behavior likely to "offend, insult, humiliate or intimidate another person" because of "race, colour or national or ethnic origin").

[118] *See* Gelber, *supra* note 115, at 8-9.

[119] *See id.* at 9-10.

accomplish this, they have both opposed and relied on international legal principles. In the 2003 decision *Hagan v. Australia*, the Committee on the Elimination of Racial Discrimination ("Committee") recommended the removal of a sign containing a racial slur from a stadium.[120] In that case, stadium management refused, and the Australian government did nothing to ensure compliance with the Committee's recommendation.[121] In a similar validation of free speech principles, the Australian High Court in *Coleman v. Power*[122] set aside a conviction under the Vagrants, Gaming and Other Offences Act[123] for the use of "any threatening, abusive, or insulting words to any person."[124] Although the majority relied primarily on the Australian constitution to enforce free speech obligations, Justice Kirby's concurrence relied on international law,[125] specifically Article 19 of the ICCPR.[126]

4. LOOKING TO THE CURRENT STATE OF HATE SPEECH REGULATION IN THE UNITED STATES TO ESTABLISH A MORE PRECISE INTERNATIONAL LEGAL STANDARD

The U.S. Supreme Court has restricted hate speech based on its content only in cases where that speech falls into a category that has been previously deemed unprotected under the First Amendment.[127] Thus, the reasoning that the Court has used to analyze government regulation of incitement,[128] fighting words,[129]

---

[120] Hagan v. Australia, U.N. GAOR, Elim. of Racial Discrim. Comm., 62d Sess., U.N. Doc. CERD/C/62/D/26/2002 (2003).

[121] Mello, *supra* note 6, at 367 (citing AAP, *UN Racism Ruling Ignored*, THE MERCURY (Austl.), July 12, 2003, at 15).

[122] *Coleman v Power* (2004) 220 CLR 1, 35 (Austl.).

[123] *Vagrants, Gaming, and Other Offences Act of 1931* (Qld.) (Austl.).

[124] *Id.* § 7(d) ("Any person who, in any public place . . . (d) [u]ses any threatening, abusive, or insulting words to any person . . . shall be liable.").

[125] *See Coleman*, 220 CLR at 82 (Kirby, J., concurring) (emphasizing the value of considering treaties and principles in international law when resolving domestic disputes).

[126] *See* ICCPR, *supra* note 12, at art. 19 (protecting the right to political expression).

[127] *See supra* notes 20–27 and accompanying texts (listing primary categories the Court recognizes).

[128] *See* Brandenburg v. Ohio, 395 U.S. 444, 447 (1969) (stating that incitement means "directed to inciting or producing imminent lawless action" and "likely to incite or produce such action").

and threats[130] underlies the Court's approach to hate speech regulation. The principle underlying this precision is speech protection. First Amendment jurisprudence has limited speech restriction primarily to speech that is closely associated with acts of violence. On the international stage, however, the right to speak is qualified in a more significant way by the non-violent effects of that speech.

Nonetheless, the international community could benefit from emulating, in part, the precision of the American system. As this Comment suggests, the international community places value on freedom from discrimination; however, the magnitude of that value varies across cases and countries. In the case of the Rwandan genocide, for example, that value increased dramatically, as discrimination led directly to mass violence. However, in cases such as *Keegstra* or *Hagan*, where the discriminatory speech is hateful but not ostensibly harmful, the line becomes difficult to draw.

To begin the process of streamlining the international system of hate speech regulation, it may be useful to incorporate a detailed imminent violence requirement.[131] Although the ICCPR prohibits speech that incites "discrimination, hostility or violence," there are few parameters around the proximity of causation.[132] In an age of instant transcontinental access, the international community must more specifically outline the role of timing between the speech and its violent or discriminatory effects through regulation.[133]

---

[129] *See* Chaplinsky v. N.H., 315 U.S. 568, 572 (1942) (providing the definition for "fighting words" as words that, when spoken, "inflict injury or tend to incite an immediate breach of the peace").

[130] *See* Planned Parenthood of the Columbia/Williamette, Inc. v. American Coalition of Life Activists, 290 F.3d 1058 (9th Cir. 2002) (stating the Court defines a true threat as a statement made with the foreseeable expectation to be understood as conveying an intent to harm); Virginia v. Black, 538 U.S. 343 (2003) (holding cross burning can be banned under the Fourth Amendment if done with *intent* to intimidate).

[131] *See Brandenburg*, 395 U.S. at 447 (requiring imminent action in its two prong test).

[132] *See* ICCPR, *supra* note 12, at art. 20(2) (relating to only "advocacy of national, racial or religious hatred").

[133] *See* MacKinnon, *supra* note 69, at 330 (noting that the *Nahimana* decision "will be carefully studied, including in cases in which incited events have yet to happen or *occur at far remove from the inciting words*") (emphasis added).

Another notion that has not yet been incorporated into international hate speech regulation involves the portion of the *Dennis* holding that the *Brandenburg* majority did not overrule. The *Dennis* decision focuses on the issue of how to treat active advocacy, or the indoctrination of members of a group to commit acts of violence, when imminence is *not* a factor.[134] The open question is, in other words: if an individual is actively advocating—as distinguished from teaching abstractly on a topic such as overthrowing the government[135]—acts of violence through hate speech, but there is no immediate plan to carry out such acts, can that individual be prosecuted for disseminating hate?

This concern is especially relevant in cases of terrorist groups whose existence is premised exclusively on the commission of acts of violence.[136] In these cases, although the U.S. Supreme Court has held against prosecution based solely on membership, evidence that a member is actively advocating violence through hate speech may be sufficient for prosecution, even if that violence is not imminent in the way that the *Brandenburg* standard prescribes. Accordingly, in the 2010 decision *Holder v. Humanitarian Law Project*,[137] the Court upheld the constitutionality of a material-support statute, which makes it a crime to "knowingly provide[] material support or resources to a foreign terrorist organization, or attempt[] or conspire[] to do so."[138] The Court concluded that *knowingly* providing support to a terrorist organization, even

[134] *See* Dennis v. U.S., 341 U.S. 494, 497 (1951) (detailing that the indictment in this case was for "wilfully [sic] and knowingly conspiring (1) to organize . . . a society, group and assembly of persons who teach and advocate the overthrow and destruction of the Government . . . by force and violence, and (2) . . . to advocate and teach the duty and necessity of overthrowing and destroying the Government . . . by force and violence," which constitute violations of Sections 2 and 3 of the Smith Act).

[135] *See id*. at 511 (citing the lower court's distinction that in order for advocacy to be considered active, it must be "reasonably and ordinarily calculated to incite persons to such action").

[136] *See* ROHAN GUNARATNA, INSIDE AL QAEDA 71–73 (2002) (detailing the extensive training of members of Al Qaeda, including instruction in the use of explosives and hand-to-hand combat as well as psychological preparedness with regard to suicide missions and the maintenance of religious zeal).

[137] Holder v. Humanitarian Law Project, 130 S. Ct. 2705 (2011).

[138] 18 U.S.C.A. § 2339B(a)(1) (2009).

without specific intent to further terrorist activity, is sufficient for prosecution.[139]

The *Holder* decision adds depth to the portion of the *Dennis* holding that *Brandenburg* did not overturn. The Court clarified that its decision in *Holder* does not constitute a reversal in precedent regarding an individual's freedom of association under the First Amendment. On the other hand, the Court emphasized that § 2339B "does *not* penalize mere association, but prohibits the act of giving foreign terrorist groups material support."[140] Further, the *Holder* decision does not prevent writing or speaking about terrorist organizations, or even adopting membership.[141] As *Holder* involved an as-applied challenge, the Court did not consider the far-reaching implications of allowing § 2339B to pass constitutional muster, thus overlooking, at least in part, the intent and imminence requirements that are central to *Brandenburg*. However, *Holder* remains an indication of flexibility and awareness of context—to which the international legal community at times affords undue weight at the expense of self-expression[142]—within the comparatively rigid *Brandenburg* paradigm.

In the international legal community, the *Holder* decision and the portion of the *Dennis* holding that this comment discusses are especially valuable when balancing speech rights—which *Brandenburg*'s intent and imminence requirements favor—with their effects on the community. In *Holder*, for example, the Court deferred to legislative findings concerning "the sensitive interests in national security and foreign affairs."[143] Accordingly, the Court found it foreseeable that materially supporting an organization whose existence is premised on terrorist activity will aid broader terrorist activity, *even if* the intent and imminence requirements of *Brandenburg* are not directly fulfilled.[144] Ultimately, the

139 Holder, *supra* note 137, at 2708–09.

140 *Id*. at 2711.

141 *Id*. at 2723.

142 *See* Ben Saul, *Speaking of Terror: Crimalising Incitement to Violence*, (2005) 28 U. NEW SOUTH WALES L.J. 868, 886 ("[E]xtending the law of incitement through new sedition offences and the power to proscribe organizations is a hasty and imprudent overreaction which inevitably criminalises valuable contributions to public discussion.").

143 Holder, *supra* note 137, at 2711.

144 *Id*.

incorporation of an imminent violence requirement and further specificity regarding speech regulation in cases where hate speech is used to indoctrinate violent groups are essential to achieving a more consistent system of hate speech regulation in international law.

5. CONCLUSION

In a time, more than any other, of instant access, the power of rhetoric is at its height. Extremist and fanatical groups disseminate hate relentlessly across all media, appealing to emotional and psychological vulnerabilities and laying an insidious groundwork of discrimination and violence. On the international stage, the issue is whether the law should respond to these messages of hate—many of them untargeted or targeted at large groups over individuals.

Europe and the UN have undertaken measures to criminalize general incitement to terrorism.[145] In 2005, for example, the Council of Europe adopted the Convention on the Prevention of Terrorism, which mandates State parties to criminalize "public provocation to commit a terrorist offence."[146] In this context, public provocation is defined as "the distribution . . . of a message to the public, with the intent to incite the commission of a terrorist offence, where such conduct . . . causes a danger that one or more such offences may be committed."[147] As a general matter, these broader criminal provisions, which would likely not pass muster under *Brandenburg*, appear to target "an environment and psychological climate conducive to criminal activity,"[148] unrelated to a specific intent to harm or offend. Though First Amendment jurisprudence in the United States does not favor content-based speech restrictions except in the most well-defined contexts,

[145] *See* Saul, *supra* note 142, at 868–69 ("Internationally, pressure to criminalise generalised incitement to terrorism has emanated from Europe and the United Nations.").

[146] Council of Europe Convention on the Prevention of Terrorism, *opened for signature* May 16, 2005, C.E.T.S. No. 6907.

[147] *Id.*

[148] Mordechai Kremnitzer and Khalid Ghanayim, *Incitement, Not Sedition*, *in* FREEDOM OF SPEECH AND INCITEMENT AGAINST DEMOCRACY 147, 197 (David Kretzmer and Francine Hazan, eds. 2000).

current political and social tensions may demand more oversight from international law.

Conversely, against this backdrop of increasing unrest and media influence, it is more important than ever to maintain the distinction between vigilance and suppression. In his book *Defending My Enemy: American Nazis, the Skokie Case, and the Risks of Freedom*, Aryeh Neier wrote, "To defend myself, I must restrain power with freedom, even if the temporary beneficiaries are the enemies of freedom."[149] This sentiment is the prevailing notion among protectors of speech rights—that hateful speech is best countered with more speech, allowing for an atmosphere of robust public debate and fueling innovation, human capital and economic development, and social change. To this end, organizations such as Human Rights Watch have determined that restricting hate speech does not necessarily further equality.[150] In addition, in countries such as Sri Lanka and South Africa, hate speech restrictions are enforced haphazardly and to the detriment of the least privileged communities.[151] However, in an era where speech is increasingly powerful, especially in the hands of groups whose sole aims are violence and destruction, it is important to continue to refine this ideal to keep pace with the changing nature of society.

International law has always recognized the inalienable rights to human dignity and freedom from discrimination; however, its efforts to protect these rights without devaluing other fundamental rights are inconsistent at best. In order to strengthen the global commitment to non-discrimination, the international community must clarify its standard for hate speech regulation. To do this, it must adopt the precision of the American speech-protective lens—a lens that declared a Birmingham parade ordinance an unconstitutional infringement on Dr. Martin Luther King, Jr.'s

---

[149] ARYEH NEIER, DEFENDING MY ENEMY: AMERICAN NAZIS, THE SKOKIE CASE, AND THE RISKS OF FREEDOM 5 (E.P. Dutton Univ. Press 1979).

[150] HUMAN RIGHTS WATCH, *'Hate Speech' and Freedom of Expression*, A HUMAN RIGHTS WATCH POLICY PAPER, Mar. 1992, at 4 (indicating that there is "little connection in practice between draconian 'hate speech' laws and the lessening of ethnic and racial violence or tension").

[151] Sandra Coliver, *Hate Speech Laws: Do They Work?*, *in* STRIKING A BALANCE: HATE SPEECH, FREE SPEECH, AND NON-DISCRIMINATION 373-74 (Sandra Coliver ed., 1992) (calling hate speech laws into question for being "vehicles for persecution of critics" and for "[compromising] the right of dissent").

right to civil protest in 1969.[152] A more coherent system of hate speech regulation in the international community will help to strengthen the relationship between free speech and freedom from discrimination for speakers and their audiences. And both sides of the conversation will be clearer.

[152] *See* Shuttlesworth v. Birmingham, 394 U.S. 147 (1969) (holding that an ordinance prohibiting any public demonstration without a permit that will be granted or withheld at the sole discretion of an official is an unconstitutional censorship of First Amendment rights).