CONTEXT MATTERS: HOW THE EUROPEAN COURT OF HUMAN RIGHTS CAN INFORM A MORE NUANCED APPROACH TO REGULATING HATE SPEECH UNDER THE FIRST AMENDMENT

ALDEN PIPER*

ABSTRACT

America has a long and ongoing history of violence aimed at minorities and historically marginalized groups. This violence has been perpetrated by state and local governments, high-ranking political officials, United States Congress, mass movements, small groups, and individuals. Violence is not created in a vacuum, and hatred towards these groups is stoked through speech that creates false narratives that present these groups as threats or stains on society. Of course, speech itself can be harmful; speech can threaten, harass, and prevent individuals from exercising their own rights. Over the last half-century, the United States Supreme Court's First Amendment jurisprudence has developed to offer hate speech absolute protection, and as a result, has prevented any regulation of hate speech despite its propensity to cause real harm. Conversely, the European Court of Human Rights ("ECtHR") deals with regulations of hate speech by employing a balancing approach that weighs the rights of the speaker with the harm created by the speech. While this approach also has drawbacks, it allows the ECtHR to evaluate context, a consideration that is lost with the U.S. Supreme Court's rigid approach. This note will seek to evaluate the merits of the ECtHR's balancing framework and will ultimately propose a new category of speech that draws on the ECtHR's balancing analysis. This proposed new category of speech endeavors to regulate hate speech under defined circumstances that produce the most harm while respecting the First Amendment's robust and essential protection of free speech.

* J.D., 2023, Boston University School of Law; B.A. History & Political Science, 2019, University of Connecticut. I would like to thank Professor Jay Wexler for research tips and guidance throughout the drafting process. I would also like to thank the *ILJ* staff for their tireless support.

CONTENTS

INTRODUCTION 367
I. LEGAL BACKGROUND 370
II. THE PROPOSAL 372
III. FORMING A NEW CATEGORY OF SPEECH 374
IV. A TWO-STEP ANALYSIS 375
V. STATIC FIRST AMENDMENT PROTECTION 379
VI. RECONCILING *R.A.V.* 381
VII. COUNTERARGUMENTS 383
CONCLUSION 387

INTRODUCTION

Freedom of speech and expression is a central tenet of America's system of self-government. America has built its free speech tradition around its citizens' right to criticize or commend a political cause and to express points of policy or opinion without censure.[1] While "Congress shall make no law. . . abridging the freedom of speech," it is up to individuals governing themselves to promote the "unhindered flow of accurate information."[2] In this way, free speech is also essential to seeking the truth.[3] The marketplace of ideas pits conflicting ideas against one another to reveal a truth and a lie or to generate a third idea that is grounded in the truth of each.[4] Freedom of speech and expression also facilitates "a central human capacity" to create and communicate through symbolic expression such as art, music, and literature, and thereby protects self-fulfillment and individual autonomy.[5] It is within this broad set of rationales in which the First Amendment's robust protection for freedom of expression is situated. Still, not all speech serves to advance these rationales, and the First Amendment treats entire categories of speech as low value and unworthy of full protection.[6] However, in contrast to approaches from other nations that guarantee freedom of speech and expression, First Amendment jurisprudence has developed to offer hate speech nearly absolute protection.

Although there is no legal definition of hate speech in the United States, it has been referenced by the Supreme Court as "[s]peech that demeans on the basis of race, ethnicity, gender, religion, age, disability or on any similar ground."[7] This definition generally aligns with the European Court of Human Rights ("ECtHR"), which defines hate speech as "all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, including intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility towards minorities, migrants and people of immigrant origin."[8] Put

---

[1] *See* ALEXANDER MEIIKKLEJOHN, FREE SPEECH: AND ITS RELATION TO SELF-GOVERNMENT 15–16 (Lawbook Exchange Ltd. 2004).

[2] *See id*.

[3] Frederick Schauer, *Free Speech, the Search for Truth, and the Problem of Collective Knowledge*, 70 SMU L. REV. 231, 235 (2017).

[4] *See* JOHN STUART MILL, ON LIBERTY 19 (Batoche Books Ltd. 2001) (1859).

[5] David A. J. Richards, *Free Speech and Obscenity Law: Toward a Moral Theory of the First Amendment*, 123 UNIV. PA. L. REV. 45, 62 (1974).

[6] *See id*. at 69–70.

[7] Matal v. Tam, 582 U.S. 218, 246 (2017).

[8] COUNCIL OF EUR., RECOMMENDATION NO. R (97) 20 ON "HATE SPEECH," at 107 (Oct. 30, 1997), https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680505d5b; *see also* ARTICLE 19, HATE SPEECH EXPLAINED — A

simply, hate speech targets individuals or groups because of who they are. Likewise, it is not surprising that such outward discrimination and hostility towards certain groups has the potential to create great harm. Writing for the majority in *Beauharnais v. Illinois*, Justice Frankfurter explained speech that discriminates against racial and religious groups "played a significant part" in leading to violence against those groups, as evidenced by the history of hate crimes and race riots in Illinois.[9] An additional example of how hate speech leads to violence is exemplified by the Rwandan genocide where the Radio-Television Libre des Milles Collines ("RTML") broadcasted hate speech encouraging acts of violence against the Tutsi minority population.[10] Following the genocide, the International Criminal Tribunal for Rwanda sentenced one of the founders of the RTML to thirty-five years in prison after finding a causal connection between RTML's broadcast of names of Tutsi individuals and those who were murdered.[11] In a more subtle example, the election of Rabbi Meir Kahane to the Knesset, who advocated for the expulsion of Arabs from Israel, led to an uptick in violence between Arab and non-Arab citizens in Israel.[12] In a more modern example, there is a "clear correlation between Trump campaign events and . . . [a] spike in hate crimes."[13]

Other than the link between hate speech and violence, advocates for hate speech regulations emphasize that hate speech laws preserve the public commitment to equal standing in society.[14] For example, Professor Waldron suggests that a well-ordered society would condemn both the leaflet that discriminates against minorities and the actions taken by individuals inspired by the leaflet's content.[15] That is because in order to have a well-ordered and just society individuals must be able to live free from fear, discrimination,

---

TOOLKIT 12 (2015), https://www.article19.org/data/files/medialibrary/38231/%27Hate-Speech%27-Explained---A-Toolkit-%282015-Edition%29.pdf [https://perma.cc/KK5V-WUYB].

[9] Beauharnais v. Illinois, 343 U.S. 250, 259 (1952) (involving case in which court upheld the conviction of petitioner under a state statute that prohibited libel against a class or citizens of a certain race, color, creed, or religion if it might cause a breach of the peace).

[10] *See* John C. Knechtle, *When to Regulate Hate Speech*, 110 DICK. L. REV. 539, 547 (2006).

[11] *See id.* at 548; Prosecutor v. Nahimana et al., Case No. ICTR-99-52-T, Summary of Judgment, at 7 (Int'l Crim. Trib. for Rwanda Nov. 28, 2007).

[12] Knechtle, *supra* note 10, at 549 (discussing passage of penal law in response to violence that punished any person who "publishes anything with the purpose of stirring up racism").

[13] Vanessa Williamson & Isabella Gelfand, *Trump and Racism: What Do the Data Say?*, BROOKINGS INST. (Aug. 14, 2019), https://www.brookings.edu/blog/fixgov/2019/08/14/trump-and-racism-what-do-the-data-say/ [https://perma.cc/PL3K-Q4PQ].

[14] *See* Jeremy Waldron, *Dignity and Defamation: The Visibility of Hate*, 123 HARV. L. REV. 1596, 1599–1600, 1627 (2010).

[15] *Id.*

and hostility and be assured that they can exercise their fundamental rights.[16] That assurance itself is a public good, and hate speech laws secure the dignity of targeted groups in society.[17] Along these lines, a number of proponents of hate speech laws suggest that victims of hate speech can actually have their own First Amendment rights curtailed and that hate speech has the power to imperil the democratic process.[18] For example, victims of hate messages feel the need to modify their behavior; they have quit jobs, foregone education, moved from their homes, and chosen to remain silent in face of hate for fear of persecution.[19]

Considering the low value nature of hate speech and the harms that it can cause, this Note argues that the Supreme Court should modify First Amendment jurisprudence to provide less protection for hate speech. Any regulation of hate speech needs to be in direct relation to its potential to cause harm and infringe on the rights of others. For example, while the European Court of Human Rights has a broad mandate to protect freedom of expression, the court still regulates hate speech by balancing the speaker's interest in freedom of expression with interest of the rights of others that are curtailed by the speech.[20] However, the downside of the ECtHR's balancing approach is its definition of hate speech is overly broad and applies generally to "all forms of expression" with a tendency to incite hatred based upon immutable characteristics or national origin.[21] This creates legal uncertainty as to what speech the ECtHR will regulate or protect and gives judges a great deal of discretion to decide cases based upon their own personal judgment as to what qualifies as hateful. Therefore, any regulation of hate speech by the Supreme Court requires a narrow definition that only applies to speech that is primarily and intentionally hateful. Specifically, this Note will suggest that the Supreme Court should adopt a two-step analysis for evaluating hate speech issues. First, the speech in question should fit a precise definition of hate speech to be considered for regulation. Second, even if the speech does amount to hate speech, the Court should then employ a balancing approach,

---

[16] *Id*. at 1596.

[17] *Id*. at 1627; *see* Beauharnais v. Illinois, 343 U.S. 250, 263 (1952).

[18] *See, e.g*., Mari J. Matsuda, *Public Response to Racist Speech*: *Considering the Victim's Story*, 87 MICH. L. REV. 2320, 2337 (1989); Katharine Gelber, *Hate Speech — Definitions & Empirical Evidence*, 32 CONST. COMMENT. 619, 625 (2017).

[19] Matsuda, *supra* note 18, at 2377–78; *see also* Richard Delgado, *Words That Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling*, 17 HARV. CIV. RTS.-CIV. LIBERTIES L. REV. 133 (1982).

[20] ARTICLE 19, *supra* note 8, at 78–80 (discussing "severity threshold" for restricting hate speech).

[21] *See, e.g*., *Factsheet — Hate Speech*, EUR. CT. HUM. RTS. (June 2022), https://www.echr.coe.int/Documents/FS_Hate_speech_ENG.pdf [https://perma.cc/EA4E-NBV4]; *Hate Speech Case Database*, FUTURE OF FREE SPEECH, https://futurefreespeech.com/hate-speech-case-database/.

like that used by the ECtHR, to evaluate if the speaker's First Amendment rights are outweighed by the victim's rights that are curtailed by the speech. While a right to be free from hate speech is not constitutionally enshrined, victims of hate speech have what can be considered their "everyday" rights infringed when they are targets of hate speech.[22] For example, most of us would recognize the right to attend a public event, work a job, or receive an education free from discrimination. The violation of these everyday rights—which would be evaluated on a case-by-case basis, would be weighed more heavily when their violation is associated with the impingement of a constitutionally protected right. These constitutional rights include the right to vote, travel (both interstate and intrastate), and the right to equal education.[23] Therefore, the rights of an individual who has to leave a voting center before casting a ballot because they are subject subjected to racist hate speech would be weighed more heavily than the rights of an individual who was compelled to leave a movie theater because of the same abuse. Ultimately, this two-step analysis is aimed at preserving the greatest number of rights while protecting the general character of the First Amendment's free speech tradition.

I. LEGAL BACKGROUND

To understand any proposal that modifies First Amendment jurisprudence in accordance with the ECtHR, it is important to first understand both the Supreme Court and ECtHR's existing legal framework. In *R.A.V. v. St. Paul*, the petitioner was convicted under a bias-motivated Minnesota Criminal Ordinance for assembling and burning a cross in the fenced yard of an African American family who lived nearby.[24] In an opinion by Justice Scalia, the Court stated that although the ordinance was aimed at "fighting words" (a category of "low-value" speech) it was unconstitutional because it singled out words that provoke violence "on the basis of race, color, creed, religion or gender," and therefore amounted to viewpoint discrimination.[25] One of

---

[22] *See* Kathryn E. Wilhem, *Freedom of Movement at a Standstill? Towards the Establishment of a Fundamental Right to Intrastate Travel,* 90 BU L. REV. 2461, 2464 (2010) (citing Johnson v. City of Cincinnati, 310 F.3d 484, 498 (6th Cir. 2002) (finding intrastate travel an "everyday right" which are rights we depend on to carry out our daily activities)).

[23] U.S. CONST. amend. XXVI; *see* Shapiro v. Thompson, 394 U.S. 618, 629 (1969) (finding a long-established right to interstate travel); Brown v. Bd. of Educ., 347 U.S. 483 (1954).

[24] R.A.V. v. City of St. Paul, 505 U.S. 377, 379 (1992).

[25] *Id.* at 391; *see* Genevieve Lakier, *The Invention of Low-Value Speech*, 128 HARV. L. REV. 2166 (2015) (describing low-value speech as speech that is offered zero or severely weakened protection under the First Amendment, and that such speech is divided into categories, including obscenity, commercial speech, true threats, child pornography and a few other categories that SCOTUS regards as having less social value and a higher potential for harm).

Scalia's main concerns was that the ordinance created asymmetry in debate; the ordinance would seemingly allow those favoring equality on the basis of race, color, creed, religion, or gender to use fighting words, while their opponents could not do the same.[26] Confusingly, the Court carved out an exception for regulations of low-value speech only where the basis for the regulation is the same reason the entire class of speech is proscribable.[27] For example, a state may permissibly choose to regulate price advertising in one industry and not another because of the risk of fraud.[28] The reason why this selective regulation is permissible is because risk of fraud is one of the reasons commercial speech is considered a low-value category of speech.[29] However, the consequence of *R.A.V.* is that hate speech cannot be regulated under current Supreme Court doctrine.[30]

Taking a divergent approach to the regulation of hate speech, the ECtHR, pursuant to the European Convention on Human Rights ("ECHR"), conducts a balancing test of the harms and benefits of the speech in question and the rights between the parties to decide if the speech should be subjected to regulation.[31] Although Article 10 of the ECHR provides for freedom of speech and expression, any guarantee of free speech must be balanced with Article 17 which states that provisions of the ECHR cannot be used to protect activity aimed at destroying rights that it guarantees.[32] Therefore, hate speech will not always enjoy the protection of Article 10 freedom of expression because it can undermine other rights guaranteed by the ECHR, such as Article 14 (prohibiting discrimination) or Article 8 (protecting the right to private and family life).[33]

As a balancing test, the ECtHR's approach to regulating hate speech is largely fact dependent. The case *Jersild v. Denmark* exemplifies the ECtHR's approach, as the court both upheld and overruled convictions based upon hate speech regulations.[34] In *Jersild*, a journalist named Jens Olaf Jersild was convicted under the Danish Penal Code for aiding and abetting the dissemination of racist remarks for broadcasting over the radio an interview with members of the Greenjackets, a xenophobic group who made

---

[26] *R.A.V.*, 505 U.S. at 391–92.

[27] *Id.* at 388.

[28] *See id.*

[29] *See* Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 44 U.S. 557 (1980) (finding a regulation aimed at commercial speech to ensure it is not misleading permissible under the First Amendment).

[30] *See R.A.V.*, 505 U.S. at 402–03 (White, J., concurring).

[31] *See, e.g.*, *Factsheet — Hate Speech*, *supra* note 21, at 19–20, 22.

[32] European Convention for the Protection of Human Rights and Fundamental Freedoms arts. 8–10, 17, Nov. 4, 1950, E.T.S. No. 5, 213 U.N.T.S. 221.

[33] *See id.* arts. 8, 10, 14.

[34] Jersild v. Denmark, 298 Eur. Ct. H.R. (ser. A) at 10–11, 15–18 (1994).

derogatory remarks above immigrants and racial minorities.[35] While the court sanctioned the convictions of the individuals who made the racist remarks, the court vacated the journalist's conviction pursuant to Article 10 of the ECHR.[36] The court stated that the broadcast pertained to an issue of public importance—racial hatred—and the conviction of Jersild would severely hamper the role of the press in disseminating material of the public interest.[37] Additionally, the court found that the objective of Jersild was to bring to light the racist attitudes of the Greenjackets, not to spread their racist message, and therefore, the objective pursued by the journalist was sufficient to outweigh any harm it may have had on the reputation or rights of others.[38] The ECtHR's balancing of harms and benefits has intuitive appeal. Under *Jersild*, the ECtHR punished bad actors and, at the same time, protected the free speech rights of the press.[39] Although not all decisions by the ECtHR have as satisfactory a result, the balancing approach provides a decision-making framework that the Supreme Court can draw on to approach hate speech issues.

II. THE PROPOSAL

Regulations based on the viewpoint of speech are particularly disfavored under the First Amendment and are typically deemed unconstitutional on their face or are subject to strict scrutiny.[40] Following *R.A.V.*, this remains true for hate speech, as the Supreme Court views hateful, derogatory, or racist speech as a viewpoint worthy of protection.[41] Additionally, the Court refuses to accept blanket prohibitions on offensive or discriminatory speech.[42] In the 2017 case *Matal v. Tam*, the Court considered a challenge to a statute that denied the registration of trademarks disparaging any person.[43] The challenge arose when Simon Tam, lead singer of the Asian-American rock group "The Slants," chose to trademark their name to reclaim the derogatory term.[44] Justice Alito, who authored the plurality opinion, refused to accept the Government's argument that the regulation was viewpoint neutral because it applied "in equal measure to any trademark that demeans or offends."[45] Instead, Alito explained that the statute amounted to unconstitutional

---

[35] *Id.* at 4, 10.

[36] *Id.* at 18–19.

[37] *Id.* at 18.

[38] *Id.* at 18–19.

[39] *See id.* at 21.

[40] *See* R.A.V. v. City of St. Paul, 505 U.S. 377, 430–31 (1992).

[41] *See* Matal v. Tam, 582 S. Ct. 218, 243 (2017) ("Giving offense is a viewpoint.").

[42] *Id.* at 246–47.

[43] *Id.* at 223.

[44] *Id.*

[45] *See id.* at 249 (Kennedy, J., concurring).

viewpoint discrimination because "giving offense is a viewpoint" and the Government has no interest in "preventing speech expressing ideas that offend."[46] Authoring a concurrence, Justice Kennedy described the danger of viewpoint discrimination as "the government . . . attempting to remove certain ideas or perspectives from a broader debate" and that the "danger is all the greater if the ideas or perspectives are ones a particular audience might think offensive."[47] However, absent from Kennedy's analysis was any defense of disparaging remarks or discriminatory messages. In fact, neither Alito nor Kennedy decided to explain the inherent value of allowing disparaging messages to go unregulated; the closest defense was Kennedy's claim that the trademark in question was not disparaging as it was intended to uplift the Asian American community.[48] In other words, context matters. However, Kennedy's concurrence makes it quite clear that the danger in regulating derogatory speech is that speech "found offensive to some portion of the public can be turned against minority and dissenting views to the detriment of all."[49] Therefore, according to Kennedy, the problem with regulating disparaging speech is not that the speech is valuable and worthy of protection, but rather that there is not an objective metric for what counts as disparaging, as well as the danger to minority viewpoints if the government is allowed to make that determination.

Still, the First Amendment offers less protection to categories of "low-value" speech based upon the speech's message and potential for harm.[50] Such categories include obscenity, commercial speech, true threats, incitement, fighting words, defamation, fraud, child pornography, and speech that is integral to criminal conduct.[51] This type of categorization presents the threat that minority viewpoints will be disregarded; reasonable minds can disagree over what is obscene, what constitutes fighting words, and what really is a true threat. However, the Supreme Court has defined and explained these categories through case law with the goal of not establishing overreaching First Amendment limitations that suppress protected speech.[52] First Amendment jurisprudence already tells us that the context of speech matters and measures can be taken to ensure that only the targeted speech

---

[46] *Id*. at 243, 246 (majority opinion).

[47] *Id*. at 250 (Kennedy, J., concurring).

[48] *Id*. at 250–51.

[49] *Id*. at 253–54.

[50] *See* R.A.V. v. City of St. Paul, 505 U.S. 377, 383, 432 (1992) (Stevens, J., concurring).

[51] *See id*. at 383–84 (majority opinion).

[52] *See, e.g.*, Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 44 U.S. 557 (1980) (establishing a multi-prong framework to determine permissible regulations of commercial speech); *see also* Miller v. California, 413 U.S. 15 (1973) (creating a definition for obscene speech based upon its "socially redeeming value" to determine if such speech is subject to First Amendment protection).

will fall under regulation.[53] Therefore, creating a low-value category of speech for hate speech necessitates a precise definition, one that accounts for the potential harm that hate speech can cause, and one that errs on the side of being underinclusive in deference to the First Amendment's robust protection of speech. As expounded upon below, the Supreme Court should regulate a new category for hate speech by balancing the harms and benefits of the speech as practiced by the ECtHR, with a concerted emphasis on the context and motivation of the speaker and the effect on the rights of the listener.

III. FORMING A NEW CATEGORY OF SPEECH

While balancing the harms and benefits of speech is largely a factual inquiry, to ensure that a new category is not overinclusive, speech that is ultimately subject to a balancing approach would first have to fall into the definition of hate speech. The current definition of hate speech suggested by the Supreme Court is speech that demeans on the basis of a protected class or group is overly broad; it focuses on the objective message the speech conveys regardless of context, when the real focus should be on the harm created by the speech and the intent of the speaker.[54] Drawing on Professor Matsuda's characterization of the worst forms of racist speech,[55] speech that falls into the category of hate speech should have the following characteristics:

(1) The message is one of inferiority based upon race, ethnicity, gender, sexual orientation, religion, or creed;

(2) The message is directed against a historically persecuted group; and,

(3) The message is purposefully persecutorial, hateful, and degrading.

The first element of the definition serves to limit the class of individuals that can avail themselves to hate speech protections to those who are members of groups based upon innate characteristics or faith-based devotion,[56] as membership in both groups is so central to an individual's culture, identity, how they perceive the world and how they are treated by others. This definition leaves out speech directed at members of political or social organizations and therefore can be seen as underinclusive. This exclusion serves to protect political speech that, despite containing hateful characteristics, is aimed at criticizing the government or other political actors. The definition also leaves out speech directed at members of a certain class, as accounted for by the second element of the definition.

---

[53] *See R.A.V.*, 505 U.S. at 426–27 (Stevens, J., concurring).

[54] *See generally* Matsuda, *supra* note 18, at 2320.

[55] *Id.* at 2357.

[56] *See id.* at 2358.

The second element focuses on the unique harm inflicted upon historically persecuted group caused by hate speech.[57] While Professor Matsuda notes that racism is the "structural subordination of a group based upon the idea of . . . inferiority,"[58] the same is true for persecution based upon sex, gender, sexual orientation, or religion. In all these contexts, hateful speech aimed at a historically persecuted group is uniquely harmful because it reinforces the systemic and societal subjugation of those groups based on immutable characteristics. This unique harm precludes socioeconomic status or class membership from protected status. Although good arguments can be made that in many instances class membership is akin to an immutable characteristic, individuals can simultaneously be members of a lower socioeconomic class and a historically dominant group. Furthermore, persecution based upon race, sex, sexual orientation, or religion transcends class; when a group is subordinated, even the wealthiest members of that group are affected.[59]

The third element considers the intent of the speaker. This element would exclude satire, misstatements, ignorant statements by speakers who do not understand the meaning of their speech, and speakers who use or reproduce hateful speech for a purpose other than to convey its hateful message.[60]

Working off this definition of hate speech (the "Proposed Definition"), the Court would then use a balancing test like that used by the ECtHR to ultimately determine whether the speech should be regulated.

IV. A TWO-STEP ANALYSIS

To determine whether hate speech should be regulated, the Supreme Court should conduct a two-step analysis to: (1) determine if the speech in question falls within the Proposed Definition; and (2) balance the harms and benefits of the speech with the rights of the victim. Case law from the ECtHR demonstrates how the Supreme Court could use a balancing approach after the speech in question is determined to fall within the Proposed Definition. The case *Perinçek v. Switzerland* involved the criminal conviction of a Turkish politician for publicly expressing the view that the mass deportations and massacres suffered by the Armenians in the Ottoman Empire in 1915 and onwards did not amount to genocide.[61] During a press conference and at a public meeting at a hotel, Doğu Perinçek claimed that the plight the

---

[57] *See id*. at 2357.

[58] *Id*. at 2358.

[59] *See id*. at 2362–63 & n.218 (noting that minority law professors, "privileged by class and education," still endure race-based harassment).

[60] *See id*. at 2357–58.

[61] Perinçek v. Switzerland, App. No. 27510/08, ¶¶ 12–16 (Oct. 15, 2015), https://hudoc.echr.coe.int/eng?i=001-158235 [https://perma.cc/S7WG-AP9V].

Armenians suffered was an "international lie" and the "Armenian problem . . . did not even exist."[62] Following these statements, the Switzerland-Armenian Association lodged a criminal complaint against the applicant for violation of Article 261 Section 4 of the Swiss Criminal Code entitled "Racial Discrimination" which deals with offenses against the public peace involving public hatred or discrimination against a person "on the grounds of their race, ethnic origin or religion."[63] A Swiss court found that the statements involved a clear "racist" motive, pointed to overwhelming historical evidence that the events in question happened, and determined the defendant was simply acting as a provocateur to suit his nationalist political agenda.[64] The defendant was found guilty, sentenced to a two-year suspended prison term, fined 3,000 francs, and ordered to pay damages to the Switzerland-Armenia Association.[65]

Following his conviction and an unsuccessful appeal, the ECtHR took up the case pursuant to the defendant's claim that the conviction violated his Article 10 rights under the ECHR.[66] Finding at the outset that the conviction clearly implicated the defendant's right to freedom of speech and expression under Article 10, the court balanced the defendant's Article 10 claim with the victim's Article 8 right to private and family life, which the court has interpreted to be triggered when there is a question of ethnic identity.[67] Here, the victim's claim was that the Armenian genocide was so central to Armenian identity that the defendant's lies undermined the dignity and identity of present-day Armenians in violation of their rights under Article 8.[68] In balancing the defendant's right under Article 10 with the harms under Article 8, the court considered that the defendant made these statements in a political capacity speaking to like-minded supporters on an issue of public concern.[69] Additionally, the court considered that the defendant did not express hatred towards the victims of the 1915 events, but rather that the strong words he used were directed towards the "imperialists."[70] Taken as a whole, the court found that the statements could not be interpreted as incitement of racial hatred, and that, as statements on a matter of public

---

[62] *Id.* ¶¶ 13–15.

[63] *Id.* ¶¶ 17–22, 32.

[64] *Id.* ¶¶ 24–26.

[65] *Id.* ¶ 22.

[66] *Id.* ¶¶ 116–17.

[67] *Id.* ¶ 180 ("The case thus entailed a conflict between two Convention rights that deserved equal protection.").

[68] *Id.* ¶ 156.

[69] *Id.* ¶ 231 ("The fact that it did not concern mainstream Swiss politics does not detract from its public interest. . . . It is the nature of political speech to be controversial and often virulent.").

[70] *Id.* ¶ 233; *see id.* ¶ 13 (referencing "England, France[,] Tsarist Russia. . . the USA and EU").

concern, they should be offered heightened protection under Article 10.[71] Therefore, the defendant's right to freedom of speech and expression under Article 10 outweighed the harm caused to the Armenian ethnic identity under Article 8.[72] Accordingly, the court found that it was not necessary to subject the defendant to a criminal conviction to protect the rights of the victims and that there was a clear violation of Perinçek's Article 10 freedom of speech and expression.[73]

The ECtHR's balancing analysis shows the importance of circumstance; the court considered the speaker's audience, his political status, the public nature of the topic, and the details of the message conveyed.[74] Although the court determined that the speech did not cause sufficient harm to override the defendant's Article 10 rights, however, it decided that the statements did at least implicate Article 8 and did not have to first define the speech as hate speech.[75] While this lack of definition produces uncertainty, it is a function of the ECtHR operating within the confines of the ECHR. Alternatively, adopting the definition of hate speech above would eliminate such uncertainty, and subsequently applying a balancing analysis similar to that employed by the ECtHR would ensure that there is still robust protection for First Amendment rights while allowing the court to safeguard the rights of others from the harm that hate speech can cause.

Returning to step one of the two-step analysis, if one applies the Proposed Definition above to the speech used in the *Perinçek* case, there is a question of whether the speech would even qualify. As to the first element, although ethnicity is clearly implicated, one would determine whether the message is one of inferiority. On one hand, the violence committed against Armenians was motivated by ethnic cleansing, an ideology based upon the notion of inferiority and dehumanization. Therefore, refuting the existence of such atrocities for political gain has the potential to promote that ideology and is at the very least an attack on Armenian identity. On the other hand, Perinçek's comments were focused on rebuking the imperialists, who he claimed fabricated the Armenian genocide to break up the Ottoman Empire and "take us hostage."[76] Therefore, the message was not necessarily one of inferiority, but was a lie told to gain political support.[77] Furthermore, Perinçek's focus on the imperialists creates a similar ambiguity as to whether the speech satisfies the second and third elements. As to the second element,

---

[71] *Id.* ¶¶ 231–41, 259.

[72] *See id.* ¶ 180.

[73] *Id.* at 115.

[74] *See id.* ¶ 66.

[75] *See id.* ¶ 228.

[76] *Id.* ¶¶ 12–13.

[77] *See id.* ¶ 22 (quoting Swiss court's conclusion that his "genocide denial" is "at least a political slogan with distinct nationalist overtone").

a question arises whether the message was directed against the Armenians, the imperialists, or both. Additionally, because Perinçek did not say anything specifically hateful towards the victims of the 1915 events or the present-day Armenian people, one would ask if the message was "persecutorial, hateful, or degrading."[78] However, given the weight of historical evidence that the Armenian genocide did occur, there is a strong argument that denying such an impactful event is purposefully hateful.

Applying the Proposed Definition to Perinçek's statements indicates that reasonable minds could disagree if the speech qualified as hate speech. Employing such a definition limits the influence that an individual judge's political or personal judgments will have on their conception of hate speech. The definition is also a way to conceptualize the harm done by the speech. In addition, the subsequent balancing test will allow a court to consider the circumstances surrounding the speech to truly evaluate the way it may have infringed on the rights of others, as well as the importance of protecting the speech in each instance. Here, the court offered Perinçek's speech heightened protection because it was arguably political speech on a matter of public concern.[79] Under the First Amendment, the speech here would also likely be considered a matter of public concern, and the harm done to a specific group would likely not be sufficient to regulate the speech.[80] Additionally, the fact that the speech was not specifically directed against the victims of the Armenian genocide or present-day Armenians would militate against the regulation of the speech in a balancing analysis.

To further explain the way this two-step analysis can safeguard First Amendment rights while accounting for the harm that hate speech causes, it is useful to examine a more straight-forward case. The 2019 case *Šimunić v. Croatia* involved a Croatian football player who repeatedly directed an expression and salute associated with the former fascist Ustashe movement towards spectators at a football match.[81] The ECtHR described the Ustashe movement as largely based upon racism and "symbolised hatred towards people of a different religion or ethnic identity . . . ."[82] Following a conviction

---

[78] Matsuda, supra note 18, at 2357.

[79] *Perinçek*, App. No. 27510/08, ¶ 231.

[80] *See* Lyrissa Lydsky, *Where's the Harm?: Free Speech and the Regulation of Lies*, 65 WASH. & LEE L. REV. 1091, 1094 (2008) (citing R.A.V. v. City of St. Paul, 505 U.S. 377, 391–92 (1992)); Lane v. Franks, 573 U.S. 228, 235 (2014) (finding "speech by citizens on matters of public concern lies at the heart of the First Amendment").

[81] Šimunić v. Croatia, App. No. 20373/17, ¶¶ 3, 44 (Jan. 29, 2019), https://hudoc.echr.coe.int/eng?i=001-189769 [https://perma.cc/6DK6-MW25]. The Ustashe was a Croatian fascist movement dedicated to independence from former Yugoslavia. The Ustashe attempted to exterminate Serbs, Jews, and Roma people to achieve a more "pure" Croatia. *See Ustasa Croatian Political Movement*, ENCYCLOPEDIA BRITANNICA, https://www.britannica.com/topic/Ustasa [https://perma.cc/GF2E-VJ4G].

[82] *Id.* ¶ 44 (discussing holding of Croatia's national courts).

for a minor offense in Croatia that resulted in a fine of 3,300 euros, the football player appealed to the ECtHR.[83] The court deferred to the Croatian court's balancing analysis between the applicant's interest in free speech and society's interest in promoting tolerance and mutual respect at public events and combatting discrimination through sport.[84] The court found that given the relatively minor fine imposed and the context in which the applicant shouted the phrase (as a famous footballer in a public setting) the Croatian court's analysis was appropriate and the applicant's Article 10 rights were outweighed by the aforementioned societal interest.[85]

The definition of hate speech would clearly apply to the applicant's speech in *Šimunić*. The speech in question was derived from a movement centered around inferiority based upon race; considering the history of the Ustashe movement, the message was clearly directed against a historically persecuted group, and, given the specificity and repetition of the expression, it is clear that it was purposefully persecutorial, hateful, or degrading. Turning to the balancing analysis, the speech in question had little value; unlike in *Perinçek*, the speech here was a personal epithet that was clearly directed against the spectators. On the other hand, obvious countervailing rights of the spectators were implicated. Individuals should have the right to attend public events free from racial animus. The cost of allowing the speech to go unregulated might deter members of historically marginalized groups from fully participating in society and attending culturally significant events, such as football games, for fear of persecution.[86] Additionally, by allowing this type of hate speech to go unregulated, current First Amendment jurisprudence burdens marginalized groups with harms to their rights in the name of First Amendment supremacy that dominant groups are not experiencing.[87] Therefore, in defined circumstances, this two step-analysis challenges the theory that the First Amendment is the most important right under the constitution[88] and acts to preserve or regulate speech based upon the weight of the harm created.

V. STATIC FIRST AMENDMENT PROTECTION

The Court has expressed reluctance to create new categories of speech, particularly those that rely on a balancing analysis.[89] Writing for the majority in *United States v. Stevens*, Chief Justice Roberts rejected the government's argument that depictions of animal cruelty should become a category of

---

[83] *Id.* ¶¶ 3, 15–19, 47.

[84] *Id.* ¶ 48.

[85] *Id.* ¶¶ 45, 48.

[86] *See* Waldron, *supra* note 14, at 1627.

[87] *See id.* at 1627, 1630–31.

[88] *See* Matsuda, *supra* note 18, at 2349.

[89] United States v. Stevens, 559 U.S. 460, 470–71 (2010).

unprotected speech.[90] One issue was that the "depictions of animal cruelty" were not closely defined and might include videos of legal activities such as hunting.[91] Roberts also criticized the government for arguing that the depictions should be subject to a simple balancing test of weighing the value of the speech against the societal costs.[92]

However, there are obvious differences between the government's proposed balancing analysis in *Stevens* versus the two-step analysis above. First, the definition of hate speech proposed in this Note limits hate speech to narrow circumstances; the Proposed Definition concerns unique harms from hate speech levied on historically marginalized groups and requires the speaker to be purposefully hateful. Second, the unique harm also distinguishes the balancing test proposed in *Stevens* from the balancing approach advocated here. In *Stevens*, the balancing test concerned the value of the speech versus the general societal costs, specifically the "social interest in order and morality."[93] This harm is somewhat amorphous, and it is difficult to imagine abridging a fundamental right in the name of an uncertain and generalized benefit to society. Conversely, the harms associated with hate speech are particular; hate speech diminishes victims' everyday rights.[94] Individuals should be able to exist in society free from hatred or discrimination and a well-ordered society would recognize their ability to do so. Additionally, in some instances hate speech will impede individuals from exercising their own constitutional rights such as the right to vote or receive an equal education. In those instances, the individual harms associated with hate speech are even more discrete and serious.[95] In this way, the proposed balancing test could be seen as an exercise where the Court aims to preserve the greatest number of rights by leaning on one side of the scale. With this contrast in mind, Roberts's rejection of the balancing test proposed in *Stevens*—finding such test as a "highly manipulable" means of identifying new categories of speech—and simultaneous refusal to "foreclose the future recognition of additional categories" makes sense.[96] Therefore, *Stevens* can be read as a rejection of a balancing approach that offers the government significant leeway to regulate speech without a showing of particularized harms, but not of the methodology of balancing as a whole.

---

[90] *Id*. at 468.

[91] *Id*. at 470, 477–78.

[92] *Id*. at 470 ("As a free-floating test for First Amendment coverage, that sentence is startling and dangerous.").

[93] *Id*. at 470 (quoting R.A.V. v. City of St. Paul, 505 U.S. 377, 383 (1992)).

[94] *See* Wilhem, *supra* note 22, at 2464.

[95] *See supra* note 23 and accompanying text.

[96] *Stevens*, 559 U.S. at 472.

VI. RECONCILING *R.A.V.*

R.A.V. v. St. Paul has had the function of prohibiting the regulation of hate speech under the First Amendment.[97] However, the decision's ambiguity—specifically the "exception to the exception"[98]—leaves open the possibility that hate speech could be regulated in accordance with *R.A.V.* without having to create a new category of speech. This would quell the concerns of Chief Justice Roberts and conform with the Court's general reluctance to create new categories of speech. However, given the dormant nature of the fighting words doctrine since *Chaplinsky*,[99] creating a new category of speech that recognizes the unique harms of hate speech seems like the better option. Still, reconciling the Court's existing case law merits a discussion.

In *R.A.V.*, all nine Justices departed from the view of the Minnesota Supreme Court that the ordinance was constitutional because it extended only to "fighting words" under *Chaplinsky*.[100] In a concurrence by Justice White, four of the Justices decided that, although the ordinance covered unprotected speech, it also covered speech that fell outside of the fighting words doctrine and was therefore unconstitutional.[101] The majority disregarded this line of reasoning and instead ruled that, even assuming the speech covered by the statute was fighting words under *Chaplinsky*, the government has an obligation of content-neutrality and therefore may not proscribe only a subset of a low-value category of speech.[102] That being said, the majority's exception to *R.A.V.*, sometimes called the "exception to the exception," said that the government may regulate within a subset of a category of speech when "the basis for the content discrimination consists entirely of the very reason the entire class of the speech is proscribable."[103] According to Scalia,

---

[97] *See R.A.V. v. City of St. Paul*, GLOB. FREEDOM EXPRESSION COLUMBIA UNIV., https://globalfreedomofexpression.columbia.edu/cases/r-v-v-city-st-paul [https://perma.cc/VA22-6JPC] ("The decision is particularly significant because it created a standard in which virtually all hate speech is protected.").

[98] *See* Elena Kagan, *The Changing Faces of First Amendment Neutrality:* R.A.V. v. St. Paul, Rust v. Sullivan, *and the Problem of Content-Based Underinclusion*, 1992 SUP. CT. REV. 29, 61–62 (1992).

[99] *See* Chaplinsky v. New Hampshire, 315 U.S. 568, 573 (1942); Burton Caine, *The Trouble with "Fighting Words:* Chaplinsky v. New Hampshire *Is a Threat to First Amendment Values and Should Be Overruled*, 88 MARQ. L. REV. 443, 547 (2004) (suggesting that "Chaplinsky has been quarantined and the fighting words doctrine rendered lame").

[100] *See* R.A.V. v. City of St. Paul, 505 U.S. 377, 381–83, 396–97 (1992) (finding that, while the ordinance extended only to fighting words under *Chaplinsky*, the ordinance was "facially unconstitutional" because it prohibits speech solely based upon the subjects the speech addresses).

[101] *Id.* at 413–14 ("The Minnesota Supreme Court erred in its application of the *Chaplinsky* fighting words test. . . .") (White, J., concurring).

[102] *See* Kagan, *supra* note 98, at 34; *R.A.V.*, 505 U.S. at 382.

[103] *R.A.V.*, 505 U.S. at 388.

a regulation in accordance with the exception is free from impermissible viewpoint discrimination.[104] The problem with this exception is that Scalia does not thoroughly explain when it could actually be satisfied.[105] "To illustrate," Scalia declares that "a State might choose to prohibit only the obscenity that is the most patently offensive *in its prurience* . . . [b]ut it may not prohibit . . . only that obscenity which includes offensive *political* messages."[106] In the context of true threats he states that

> the Federal Government can criminalize only those threats of violence that are directed against the President, since the reasons why threats of violence are outside the First Amendment (protecting individuals from the fear of violence, from the disruption that fear engenders, and from the possibility that the threatened violence will occur) have special force when applied to the person of the President.[107]

Therefore, the first step in the analysis under *R.A.V.* is to determine the reason why a category of speech is considered low-value and is offered less than full First Amendment protection.

In *Chaplinsky*, the Court described fighting words as "those which by their very utterance inflict injury or tend to incite an immediate breach of the peace."[108] The Court went on to note that "such utterances are no essential part of any exposition of ideas, and are of such slight social value as a step to truth that any benefit . . . is clearly outweighed by the social interest in order and morality."[109] The Court seems to justify the regulation by pointing to the inability of fighting words to contribute to Mill's marketplace of ideas. Central to Mill's conception is that low-value or false speech should not be restricted because, through competition with all other speech, the truth will prevail.[110] However, the Court takes the view that fighting words can be regulated because they are of such low-value that any contribution to the marketplace is clearly outweighed by the harm it causes.[111] Therefore, fighting words that are regulated because they fail to pass the value-versus-harm-caused balancing test set forth in *Chaplinsky* fall within the *R.A.V.* exception. Ironically, following solely the Court's own language brings us back to a balancing approach that Chief Justice Roberts and the current Court

---

[104] *Id.*

[105] *See id.* at 423–24 (characterizing exception and explanation as "opaque") (Stevens, J., concurring).

[106] *Id.* at 388 (majority opinion) (emphasis in original).

[107] *Id.* (internal citations omitted).

[108] Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942) (citing ZECHARIAH CHAFEE, JR., FREE SPEECH IN THE UNITED STATES 149 (1941)).

[109] *Id.* (citing CHAFEE, *supra* note 108, at 150).

[110] *See* MILL, *supra* note 4, at 19.

[111] *See Chaplinsky*, 315 U.S. at 572.

seems to reject.[112] Additionally, following the *R.A.V.* "exception to the exception" in the fighting words context would undermine the main holding of *R.A.V.* because the same speech that could not be regulated in *R.A.V.* could now be permissibly regulated simply because of alternative justification. Therefore, a narrower definition of hate speech than that used in the Minnesota ordinance may be needed to pass such a balancing test. The Minnesota ordinance prohibited the display of a symbol that an individual "knows or has reason to know arouses anger, alarm, or resentment in others on the basis of race, color, creed, religion or gender. . . ."[113] Messages or displays that simply cause anger or resentment based on an immutable characteristic could encompass political or activist displays. For example, a protestor holding a sign criticizing affirmative action, Title IX programming, or the state of Israel could conceivably fall within this ordinance because each of those displays has the propensity to cause anger because of a viewer's race, gender, or religion. In contrast, the definition of hate speech proposed by this Note above relies on a message of inferiority based upon immutable characteristics that is purposefully "persecutorial, hateful, or degrading."[114] An ordinance written on similar terms would not likely apply to speech that has value; instead, it would regulate speech that impedes an individual's rights, not merely cause anger, alarm, or resentment.

VII. COUNTERARGUMENTS

One of the most common arguments against any hate speech regulation is that the act of regulating unprotected speech will stifle protected speech.[115] This chilling effect occurs because speech regulations are vague or overly broad and prevent speakers from knowing what speech is covered by the regulations.[116] In turn, individuals might self-censor otherwise permissible speech for fear of liability.[117] As Chemerinsky notes, courts across the country are extremely wary of the chilling effect of hate speech regulations and have struck down numerous hate speech codes that have been adopted by public universities as overly broad.[118] For example, courts have taken issue with regulations that ban language that "stigmatizes" based on "race,

---

[112] *See* United States v. Stevens, 559 U.S. 460, 469 (2010).

[113] R.A.V. v. City of St. Paul, 505 U.S. 377, 379 (1992)

[114] Matsuda, *supra* note 18, at 2357.

[115] *See generally* Erwin Chemerinsky, *The Challenge of Free Speech on Campus*, 61 HOW. L.J. 585 (2018).

[116] *Id.* at 595.

[117] *Id.*

[118] *Id.* at 592–93 (discussing UWM Post, Inc. v. Bd. of Regents of Univ. of Wis. Sys., 774 F. Supp. 1163 (E.D. Wis. 1991) and IOTA XI Chapter of Sigma Chi Fraternity v. George Mason Univ., 993 F.2d 386 (4th Cir. 1993)).

religion, sex, sexual orientation, [or] creed,"[119] or codes that prohibit speech that is "insulting" or "demeaning."[120] Indeed, there is much force to the argument that because there is not a homogenous view as to what should be considered "insulting" or "demeaning" any regulation of speech on those grounds would be inherently political.[121]

The two-step analysis proposed in this Note suffers from similar uncertainty. While the speech in question must first fall within the Proposed Definition of hate speech, it will then be subject to a balancing test concerned with the circumstances of each case and preserving the greatest number of rights. Therefore, although the two-step analysis is specifically designed to guard against the regulation of speech that is innocuous, academic, political, comedic, or simply based on a lack of knowledge, it will never offer complete guidance for speakers to know exactly what language it covers.

However, the proposed definition of hate speech does require the language to be a message of inferiority directed against a historically subjugated group that is purposefully "persecutorial, hateful, and degrading."[122] The *mens rea* requirement gives speakers notice that so long as they do not speak with an intent to cause harm their speech will not fall within the Proposed Definition. Furthermore, the subsequent balancing test introduces a seriousness requirement and is much more rigorous than a hate speech code that punishes individuals for levying insults at one another. Under the balancing approach, speech can only be regulated if the harm it inflicts is serious enough that the rights of the speaker are outweighed by the rights of the listener. Therefore, while there will never be complete certainty for speakers as to what may or may not be spoken, the specificity of the proposed definition, the *mens rea* requirement, and the seriousness requirement provide levels of notice that are lacking in other hate speech codes.[123]

Another powerful argument against regulating hate speech is that such regulations will not reduce discrimination and may even aggravate discriminatory behavior.[124] Former president of the American Civil Liberties Union, Nadine Strossen, cites empirical evidence from Great Britain that shows the introduction of racial defamation laws had no impact on the activity of neo-Nazi groups over twenty-five years after their implementation.[125] However, Strossen's criticism of hate speech regulations on these grounds depends on her belief that the goal of hate speech regulation

---

[119] *Id*. at 591 (quoting Doe v. Univ. of Mich., 721 F. Supp. 852, 856 (E.D. Mich. 1989)).

[120] *Id*. at 595–96.

[121] *See id*. at 596.

[122] Matsuda, *supra* note 18, at 2357.

[123] *See* Chemerinsky, *supra* note 115, at 585.

[124] Nadine Strossen, *Regulating Racist Speech on Campus: A Modest Proposal*, 1990 DUKE L.J 484, 556 (1990).

[125] *Id*. at 554–55.

is to "combat" racism and discrimination.[126] While this is a laudable social objective, Strossen rightly states that "racist speech is one symptom of the pervasive problem of racism, and this underlying problem will not be solved by banning one of its symptoms."[127]

Alternatively, rather than providing an antidote to an endemic social ill, the proposed two-step analysis is concerned with vindicating the rights of the victim in a particular instance. As Waldron contends, a well-ordered society would not sit idly by in the face of hateful displays that diminish individual dignity.[128] Furthermore, legal regulation of hate speech provides some solace to victims that society at large recognizes the harm of hate speech.

Even if one operates under the belief that the primary goal of hate speech regulation is to reduce discrimination, it is possible that regulations will help achieve that goal. Strossen contends that censoring hate speech will stifle essential intergroup dialogue around issues of discrimination that would otherwise reduce such discrimination.[129] Therefore, she believes that, although hate speech regulations may prevent racist or other discriminatory remarks from being spoken aloud, such regulations perpetuate discriminatory behavior by curtailing conversation about why those forms of behavior or communication are harmful or wrong.[130]

However, hate speech is based on patterns of discrimination or stereotypes that have historical roots and have been perpetuated over time. For example, racist theories such as eugenics are still relied upon by perpetrators of hate speech today, and the same stereotypes used against Asian Americans during the time of the Chinese Exclusion Act have resurfaced throughout the COVID-19 pandemic.[131] Therefore, regulations may reduce hate speech in the aggregate by preventing some hate speech from being spoken aloud and

---

[126] *See id*. at 556–61.

[127] *Id*. at 554; *see also id*. at 559 ("Advocates of hate speech regulations do not seem to realize that their own attempts to suppress speech increase public interest in the ideas they are trying to stamp out.").

[128] Waldron, *supra* note 14, at 1610–13 (discussing that "respect for human dignity. . . is a crucial foundation of basic rights and equality").

[129] *See* Strossen, *supra* note 124, at 561.

[130] *See id*.

[131] *See* Peter J. Breckheimer, Note, *A Haven for Hate: The Foreign and Domestic Implications of Protecting Internet Hate Speech Under the First Amendment*, 75 S. CAL. L. REV. 1493, 1496 (2002); Amanda Mull, *Americans Can't Escape Long Disproven Body Stereotypes*, ATLANTIC (Nov. 6, 2018), https://www.theatlantic.com/health/archive/2018/11/body-stereotypes-personality-debunked-eugenics/575041 [https://perma.cc/EM9V-86HB]; Helier Cheung et al., *Coronavirus: What Attacks on Asians Reveal About American Identity*, BBC NEWS (May 27, 2020), https://www.bbc.com/news/world-us-canada-52714804 [https://perma.cc/6AQ8-34SJ]; Angela R. Gover, Shannon B. Harper & Lynn Langton, *Anti-Asian Hate Crime During the COVID-19 Pandemic: Exploring the Reproduction of Inequality*, 45 AM. J. CRIM. JUST. 647 (2020).

limiting its subsequent reproduction by others.

Furthermore, there is a distinction between the hate speech that Strossen refers to as "failures of sensitivity" and the hate speech that would be regulated by the proposed two-step analysis.[132] Harmful speech that emerges from a lack of knowledge or understanding, i.e., a failure of sensitivity, is exactly the type of speech that should be openly discussed. In those instances, dialogue can promote greater tolerance and educate speakers about the negative impact that speech can have. However, it is hard to imagine that an open dialogue centered around education and tolerance will have similar beneficial effects with a speaker whose aim is to be intentionally hateful.

A difficult issue raised by Strossen and other opponents of hate speech codes is the contention that speech regulations are often used against minority groups they are trying to protect.[133] Among other examples, Strossen cites a now-rescinded hate speech rule from the University of Michigan that resulted in over twenty white students charging black students with hate speech.[134] In the only two instances where the rule was actually enforced, the rule sanctioned two black students.[135] Additionally, outside of the university context, the first individuals prosecuted under the British Race Relations Act of 1965, an act designed to stifle the Neo-Nazi National Front, were black power leaders who used racist language while speaking out against discrimination and violence inflicted against black people.[136] The fact that hate speech regulations vest so much power in university officials, judges, and prosecutors to determine who will fall within their purview is a great concern. These officials' own biases will likely infect the law enforcement process and with politicized judgments about who the law should apply to. Although the definition of hate speech proposed in this Note is not a complete safeguard from this type of selective enforcement, the emphasis on historically persecuted groups would tailor enforcement to those groups where hate speech is likely to inflict unique harm. As previously stated, hate speech inflicted on historically persecuted groups creates a unique harm by reinforcing the systemic and structural subordination of a group. Alternatively, hate speech levied upon historically dominant groups is less likely to impact the rights of the victim because they do not carry with them a history of individual or state-sponsored violence.[137]

Strossen and Chemerinsky cite examples where minority students were

---

[132] Strossen, *supra* note 124, at 561.

[133] *Id*. at 556.

[134] *Id*. at 557.

[135] *Id*. at 557.

[136] *Id*. at 556; *see* Michel Rosenfeld, *Hate Speech in Constitutional Jurisprudence: A Comparative Analysis*, 24 CARDOZO L. REV 1523, 1546–47 (2003) (one member of the Black Liberation Movement was convicted under the statute for asserting that "whites are vicious and nasty people" and recounting how he saw "white savages" kicking black women).

[137] *See* Matsuda, *supra* note 18, at 2362.

punished for making discriminatory comments as instances where hate speech regulations have backfired by punishing those they are intended to protect. Both scholars cite an example offered by Professor Henry Louis Gates, where under the University of Michigan hate speech rule, the only full disciplinary hearing conducted involved a black student who said that "homosexuality was an illness" and he was developing a project to "move homosexuals towards heterosexuality."[138] Additionally, Strossen cites a case at the University of Connecticut where an Asian-American student was sanctioned under a university policy for a homophobic remark.[139] Although these examples may be cited for the proposition that hate speech regulations might be disproportionately used against minorities, they seem to misconstrue the goal of hate speech regulation. Hate speech regulations should not preclude punishing a member of one historically persecuted group for saying purposefully hateful comments about a member of a separate historically persecuted group. They are not implemented to offer minority groups absolute protection both as a speaker and listener. Chemerinsky and Strossen's examples seem to suggest that these regulations backfired because they were used against minorities but fail to discuss that the students in both instances were alleged to have said abhorrent homophobic statements. Ultimately, the proposed two-step analysis moves away from conceptualizing hate speech regulations as measures designed to only protect or punish certain groups, but as measures designed to preserve the greatest number of rights in each instance.

CONCLUSION

Freedom of thought and expression is essential to creating a society that values open dialogue and seeking the truth.[140] Even those beliefs that are found nearly universally repugnant, if held to the light through communication and robust debate, can play a role in the truth-seeking process. Authors, comedians, painters, politicians, or protestors might want to use these offensive or harmful ideas to draw attention to an issue or to highlight absurdity. The Supreme Court recognizes that as human beings these actors should have the autonomy to communicate through symbolic expression without the sanction of the state.[141] When Paul Robert Cohen was

---

[138] *See* Chemerinsky, *supra* note 115, at 597; Strossen, *supra* note 124, at 527–28 n.211; HENRY LOUIS GATES ET AL., SPEAKING OF RACE, SPEAKING OF SEX: HATE SPEECH, CIVIL RIGHTS, AND CIVIL LIBERTIES 45 (1995).

[139] *See* Strossen, *supra* note 124, at 558 (discussing Wu v. Univ. of Conn., No. Civ. H89-649 (D. Conn. Jan. 19, 1990)).

[140] *See* MILL, *supra* note 4, at 21–22.

[141] Richards, *supra* note 5, at 62; *see* United States v. O'Brien, 391 U.S. 367 (1968) (finding that the government may only regulate symbolic expression if the law satisfies a four-part test and, among other requirements, is the least restrictive means of regulation).

convicted of disturbing the peace for wearing a jacket in a Los Angeles courthouse bearing the words "Fuck the Draft," Justice Harlan wrote the following: "[W]e cannot indulge the facile assumption that one can forbid particular words without also running a substantial risk of suppressing ideas in the process. Indeed, governments might soon seize upon the censorship of particular words as a convenient guise for banning the expression of unpopular views."[142] As Justice Harlan and Chemerinsky state, it is naive to believe you can suppress words without suppressing ideas.[143] However, society is capable of distinguishing between speech that is socially or politically relevant and speech whose only utility is to inflict harm.

The new category for hate speech proposed in this Note would allow courts to analyze the circumstances surrounding the speech before they are bound to determine if a First Amendment violation has occurred. This is exactly how the Supreme Court operates when considering if speech falls into other "low-value" categories.[144] In the context of obscenity, material could be viewed as obscene in one community and not another, and context would dictate whether obscene material had artistic or political value. Additionally, the two-step analysis accounts for traditional critiques of hate speech regulation, i.e., the dangers of censorship and selective enforcement. The Proposed Definition for hate speech is narrow and requires an intent to cause harm. Furthermore, even if speech does meet the definition of hate speech, the Court must then balance the rights of the speaker with the rights of the listener to determine if the speech can be suppressed. This balancing approach, modeled after the ECtHR, refocuses the Court from simply considering if speech was impermissibly suppressed to analyzing what rights are at stake and allows the Court to protect the greatest number of rights. One of the primary functions of the Constitution is to protect the rights of American citizens; therefore, a Court that is ostensibly concerned with promoting open and free debate, tolerance, and seeking the truth through robust First Amendment protection should, at its core, be concerned with protecting rights.

Ultimately, the proposed category for hate speech creates a high bar for enforcement. The function of the category is not to promote a certain ideology or to substitute more meaningful approaches to combatting hatred and discrimination.[145] The speech captured by the proposed category is only

---

[142] Cohen v. California, 403 U.S. 15, 26 (1971).

[143] *See id.*; Chemerinsky, *supra* note 115, at 599.

[144] For example, in the case of obscenity, the Court applies the *Miller* test, stating material is obscene if: (1) an average person applying contemporary community standards would find the material appeals to the prurient interest; (2) the material is patently offensive; and (3) whether the work taken as a whole, lacks serious literary, artistic, political, or scientific value. Miller v. California, 413 U.S. 15, 24 (1973).

[145] *See* Strossen, *supra* note 124, at 561.

the speech that is the most vitriolic and capable of inflicting unique harm.[146] The category simply gives the Court the ability to protect the rights of the listener if they outweigh the right of the speaker to communicate in an intentionally hateful manner. This level of enforcement would not stifle America's commitment to free speech; it would uphold the country's commitment to liberty and equality under the law.

[146] *See supra* notes 57–58.