

HASHTAG HATE: THE NEED FOR REGULATING MALIGNANT RHETORIC ONLINE

Stephanie J. Beach*

| | |
|---|-----|
| INTRODUCTION..... | 129 |
| I. THE HISTORY OF HATE SPEECH AND FIRST AMENDMENT | |
| JURISPRUDENCE | 131 |
| A. Incitement to Imminent Lawless Action..... | 132 |
| B. Fighting Words | 136 |
| C. True Threats | 139 |
| II. HATE SPEECH ONLINE AND HATE CRIMES IN PERSON..... | 140 |
| III. A PRESCRIPTION FOR LEGAL AND REGULATORY REFORM | 150 |
| A. Private Action | 154 |
| B. Governmental Action..... | 155 |
| C. Global Action..... | 157 |
| D. Judicial Action | 159 |
| E. Oversight Review Board..... | 160 |
| CONCLUSION | 162 |

INTRODUCTION

On Saturday, October 27, 2018, a 46-year-old man armed with an AR-15 semi-automatic rifle and three handguns entered the unlocked front door of the Tree of Life Congregation in Pittsburgh, Pennsylvania and systematically murdered eleven congregants, wounded two others, and injured four responding law enforcement officers.¹ Five minutes before police were notified of the shooting, the gunman posted on his social media account, “I can’t sit by and watch my people get slaughtered . . . Screw your optics, I’m going in.”² The shooter littered his Gab page with anti-Semitic

*. J.D., Seton Hall University School of Law; M.S., Columbia University Graduate School of Journalism; B.A., New York University. The Author wishes to thank her family for acting as a sounding board for the ideas set forth in this Article, and especially her grandmother, whose endurance through extreme hatred served as the inspiration for this Article.

1. Campbell Robertson, Christopher Mele & Sabrina Tavernise, *11 Killed in Pittsburgh Synagogue Massacre; Suspect Charged With 29 Counts*, N.Y. TIMES (Oct. 27, 2018), <https://www.nytimes.com/2018/10/27/us/active-shooter-pittsburgh-synagogue-shooting.html>.

2. Dakin Andone, Jason Hanna, Joe Sterling, & Paul P. Murphy, *Hate Crime Charges Filed in Pittsburgh Synagogue Shooting That Left 11 Dead*, CNN (Oct. 29, 2018), <https://www.cnn.com/2018/10/27/us/pittsburgh-synagogue-active-shooter/>.

jargon, including a bio that read, “jews are the children of satan.”³ Gab, a social media network that promotes virtually total freedom of speech, is an alternative to Facebook and Twitter, and has found devotees amongst a more conservative user base.⁴

The vast majority of Americans have accounts on at least one social media platform.⁵ And these social media users report increasing online harassment based on inherent personal qualities.⁶ There are a problematically large number of websites, social networking pages, and forums that are increasingly used to promote malignant rhetoric.⁷ Sixty-two percent of Americans surveyed viewed online harassment “as a major problem,” and seventy-nine percent feel that online services “have a duty to step in when harassment occurs on their platforms.”⁸ Because what commences as hate speech on social media often leads to hate crimes in person.⁹

This Article will address the conundrum of the rising plight of hate speech on social media platforms and propose a framework for combatting the upsurge of bias online. Part I will delve into the history of the law governing hate speech as crafted by United States Supreme Court

3. *11 Dead, 6 Injured in Pittsburgh Synagogue Shooting*, JEWISH TELEGRAPHIC AGENCY (Oct. 27, 2018), <https://www.jta.org/2018/10/27/top-headlines/least-4-reported-dead-pittsburgh-synagogue-shooting>.

4. Alex Kantrowitz, *This New Social Network Promises Almost Total Free Speech to Its Users*, BUZZFEED NEWS (Sept. 9, 2016), <https://www.buzzfeednews.com/article/alexkantrowitz/new-social-network-gab-growing-fast-free-speech>.

5. *Social Media Fact Sheet*, PEW RESEARCH CTR. (June 12, 2019), <http://www.pewinternet.org/fact-sheet/social-media/>.

6. *Online Hate and Harassment: The American Experience*, ANTI-DEFAMATION LEAGUE, <https://www.adl.org/onlineharassment> (last visited Dec. 11, 2019) [hereinafter *Online Hate and Harassment*].

7. *Social Media Must Do More to Thwart Subculture of Hate Fueling Lone Wolf Terrorism - Simon Wiesenthal Center Debuts 2012 Digital Hate Report*, SIMON WIESENTHAL CTR. (Mar. 27, 2012), <http://www.wiesenthal.com/about/news/social-media-must-do-more-to.html> [hereinafter *Social Media Must Do More*].

8. Maeve Duggan, *Online Harassment 2017*, PEW RESEARCH CTR. (July 11, 2017), http://www.pewinternet.org/2017/07/11/online-harassment-2017/?utm_source=AdaptiveMailer&utm_medium=email&utm_campaign=17-7-11%20Online%20Harassment&org=982&lvl=100&ite=1409&lea=303148&ctr=0&par=1&trk=.

9. A hate crime is “a traditional offense like murder, arson, or vandalism with an added element of bias.” *Hate Crimes*, FBI, <https://www.fbi.gov/investigate/civil-rights/hate-crimes> (last visited Dec. 11, 2019). “The FBI has defined a hate crime as a ‘criminal offense against a person or property motivated in whole or in part by an offender’s bias against a race, religion, disability, sexual orientation, ethnicity, gender, or gender identity.’” *Id.* See Rachel Hatzipanagos, *How Online Hate Turns Into Real-Life Violence*, WASH. POST (Nov. 30, 2018), https://www.washingtonpost.com/nation/2018/11/30/how-online-hate-speech-is-fueling-real-life-violence/?utm_term=.902d0038dfcc (explaining how “online messages can turn into real-life violence”).

jurisprudence.¹⁰ Part II will explore the expansion of hate speech online.¹¹ Part III explains the need for change and sets forth an agenda for legal and regulatory reform.¹² Finally, this Article concludes by briefly summarizing this paradox.¹³

I. THE HISTORY OF HATE SPEECH AND FIRST AMENDMENT JURISPRUDENCE

The First Amendment of the Bill of Rights provides that: “Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for a redress of grievances.”¹⁴ The Fourteenth Amendment extended this freedom to the individual states.¹⁵

The First Amendment does not protect all speech.¹⁶ What it does protect is speech that is integral to advancing the marketplace of ideas.¹⁷ In his dissent in *Abrams v. United States*, Justice Oliver Wendell Holmes, Jr. wrote:

But when men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of their own conduct 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, and that truth is the only ground upon which their wishes safely can be carried out. That at any rate is the theory of our Constitution. It is an experiment, as all life is an experiment.¹⁸

10. *See infra* Part I.

11. *See infra* Part II.

12. *See infra* Part III.

13. *See infra* pp. 34–35.

14. U.S. CONST. amend. I.

15. U.S. CONST. amend. XIV; *Incorporation Doctrine*, LEGAL INFO. INST., https://www.law.cornell.edu/wex/incorporation_doctrine (last visited Dec. 11, 2019).

16. *Chaplinsky v. New Hampshire*, 315 U.S. 568, 571–72 (1942) (“[I]t is well understood that the right of free speech is not absolute at all times and under all circumstances. 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.” (footnotes omitted)).

17. “The marketplace of ideas refers to the belief that the test of the truth or acceptance of ideas depends on their competition with one another and not on the opinion of a censor, whether one provided by the government or by some other authority.” David Schultz, *Marketplace of Ideas*, THE FIRST AMENDMENT ENCYCLOPEDIA (last updated June 2017), <https://www.mtsu.edu/first-amendment/article/999/marketplace-of-ideas>.

18. *Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

He continued, explaining that “we should be eternally vigilant against attempts to check the expression of opinions that we loathe and believe to be fraught with death, unless they so imminently threaten immediate interference with the lawful and pressing purposes of the law that an immediate check is required to save the country.”¹⁹

Hate speech is speech that is perceived as harmful and offensive to racial or religious minorities or other historically disempowered groups.²⁰ It is “used as [a] weapon[] to ambush, terrorize, wound, humiliate, and degrade.”²¹ The United States Supreme Court has rendered numerous decisions on hate speech.²² These decisions constructed the constitutional framework for determining what speech is permissible and what speech can be criminalized.

A. Incitement to Imminent Lawless Action

Hate speech can only be criminalized if it is an “incitement to imminent lawless action,”²³ or it “communicate[s] a serious expression of an intent to commit an act of unlawful violence to a particular individual or group of individuals.”²⁴

The Supreme Court developed the “clear and present danger” test in a 1919 case.²⁵ In *Schenck v. United States*, Justice Holmes wrote that “[t]he question in every case is whether the words used are used in such circumstances and are of such a nature as to create a clear and present danger that they will bring about the substantive evils that Congress has a right to prevent.”²⁶ This test instructed courts to look to the probable results of the speech, irrespective of whether the results actually occurred.²⁷ This permitted the Court to uphold a conviction for two individuals who published a newspaper criticizing World War I.²⁸ There, Justice Holmes again wrote that “it is impossible to say that it might not have been found that the circulation of the paper was in quarters where a little breath would be enough to kindle

19. *Id.*

20. MARI J. MATSUDA ET AL., WORDS THAT WOUND: CRITICAL RACE THEORY, ASSAULTIVE SPEECH, AND THE FIRST AMENDMENT 1 (1993).

21. *Id.*

22. *See infra* Parts I.A–I.C.

23. *Brandenburg v. Ohio*, 395 U.S. 444, 449 (1969).

24. *Virginia v. Black*, 538 U.S. 343, 359 (2003) (citing *Watts v. United States*, 394 U.S. 705, 708 (1969)).

25. *Schenck v. United States*, 249 U.S. 47, 52 (1919).

26. *Id.*

27. *Id.*

28. *Frohwerk v. United States*, 249 U.S. 204, 210 (1919).

a flame and that the fact was known and relied upon by those who sent the paper out.”²⁹

But during the same year, in his dissent in *Abrams*, Justice Holmes explained that there is no doubt that “the United States constitutionally may punish speech that produces or is intended to produce a clear and imminent danger that it will bring about forthwith certain substantive evils that the United States constitutionally may seek to prevent.”³⁰ He opined that “we should be eternally vigilant against attempts to check the expression of opinions that we loathe and believe to be fraught with death, unless they so imminently threaten immediate interference with the lawful and pressing purposes of the law that an immediate check is required to save the country.”³¹ Thus, he added the mandate of immediacy by stating that “[i]t is only the present danger of immediate evil or an intent to bring it about that warrants Congress in setting a limit to the expression of opinion where private rights are not concerned.”³²

Yet in a 1917 case before the United States Court of Appeals for the Second Circuit, Judge Rodgers suggested a different approach, opining that instead of pondering the probable results of the speech, courts should only consider the direct results.³³ In a 1951 case, the Court adopted a previous statement by then Chief Judge Learned Hand, who stated that courts must “ask whether the gravity of the ‘evil,’ discounted by its improbability, justifies such invasion of free speech as is necessary to avoid the danger.”³⁴ This modified the Court’s original “clear and present danger” test.³⁵ Now, a conviction could still stand if the harm was significant, even if the risk was diminutive. But six years later, the Court distinguished *Dennis* by holding that there was a “distinction between advocacy of abstract doctrine and advocacy directed at promoting unlawful action.”³⁶

One year later, in 1952, in an opinion penned by Justice Frankfurter, the Court upheld an Illinois criminal libel law that made it illegal to publish or exhibit any writing or picture portraying the “depravity, criminality, unchastity, or lack of virtue of a class of citizens of any race, color, creed or

29. *Id.* at 209.

30. *Abrams v. United States*, 250 U.S. 616, 627 (1919) (Holmes, J., dissenting).

31. *Id.* at 630.

32. *Id.* at 628.

33. *Masses Pub. Co. v. Patten*, 246 F. 24, 38 (2d Cir. 1917).

34. *Dennis v. United States*, 341 U.S. 494, 510 (1951) (quoting *United States v. Dennis*, 183 F.2d 201, 212 (2d Cir. 1950)).

35. *Schenck v. United States*, 249 U.S. 47, 52 (1919).

36. *Yates v. United States*, 354 U.S. 298, 318 (1957).

religion.”³⁷ He explained that while “the possibility of abuse is a poor reason for denying Illinois the power to adopt measures against criminal libels sanctioned by centuries of Anglo-American law,” the statute in question was not impermissibly broad or vague.³⁸ The Court expounded that: “Libelous utterances not being within the area of constitutionally protected speech, it is unnecessary, either for us or for the State courts, to consider the issues behind the phrase ‘clear and present danger.’”³⁹ Thus, this provided that the First Amendment may not protect hate speech that violates libel law.

It was not until 1969 in *Brandenburg v. Ohio*—a case involving a Ku Klux Klan (KKK) leader who was convicted under an Ohio statute for “advocat[ing] . . . the duty, necessity, or propriety of crime, sabotage, violence, or unlawful methods of terrorism as a means of accomplishing industrial or political reform” as well as “voluntarily assembl[ing] with any society, group, or assemblage of persons formed to teach or advocate the doctrines of criminal syndicalism”—that the Court refined its incitement test.⁴⁰ There, the Court held that

the 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.⁴¹

This concept can be broken down into four requisite elements: (1) intent to incite lawless action; (2) intent to incite lawless action imminently; (3) likelihood of inciting lawless action; and (4) likelihood of inciting lawless action imminently. The Court found that a “statute which fails to draw this distinction impermissibly intrudes upon the freedoms guaranteed by the First and Fourteenth Amendments” by “sweep[ing] within its condemnation speech which our Constitution has immunized from governmental control.”⁴² The concurrence distinguished that the “line between what is permissible and not subject to control and what may be made impermissible and subject to regulation is the line between ideas and overt acts.”⁴³

37. *Beauharnais v. Illinois*, 343 U.S. 250, 251 (1952) (citation omitted).

38. *Id.* at 263–64.

39. *Id.* at 266.

40. *Brandenburg v. Ohio*, 395 U.S. 444, 444–45 (1969) (alterations in original).

41. *Id.* at 447 (footnote omitted).

42. *Id.* at 448.

43. *Id.* at 456 (Douglas, J., concurring).

In *Collin v. Smith*, the National Socialist Party of America planned a march through the streets of the Village of Skokie, Illinois, a community that thousands of Holocaust survivors called home.⁴⁴ The Village issued an ordinance that mandated (as a prerequisite for obtaining a permit to march) that the group “will not portray criminality, depravity or lack of virtue in, or incite violence, hatred, abuse or hostility toward a person or group of persons by reason of reference to religious, racial, ethnic, national or regional affiliation.”⁴⁵ Another ordinance prohibited “the dissemination of any materials within the Village of Skokie which promotes and incites hatred against persons by reason of their race, national origin, or religion, and is intended to do so.”⁴⁶ The Seventh Circuit Court of Appeals stated that its “task . . . is to decide whether the First Amendment protects the activity in which [the National Socialist Party of America] wish[es] to engage, not to render moral judgment on their views or tactics.”⁴⁷ The Village asserted “that *Beauharnais* implicitly sanctions prohibiting the use of First Amendment rights to invoke racial or religious hatred *even without reference to fears of violence*.”⁴⁸ The Seventh Circuit found this argument insufficient.⁴⁹ Instead, the court found that the “government has no power to restrict expression because of its message, its ideas, its subject matter, or its content.”⁵⁰ In quoting *N.Y. Times Co. v. Sullivan*,⁵¹ the Seventh Circuit impressed the “profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open.”⁵² The Court did concede that “[i]t is, after all, in part the fact that our constitutional system protects minorities unpopular at a particular time or place from governmental harassment and intimidation, that distinguishes life in this country from life under the Third Reich.”⁵³ As such, hate speech can either be viewed as fighting words or true threats.

44. *Collin v. Smith*, 578 F.2d 1197, 1199 (7th Cir. 1978).

45. *Id.* (citation omitted).

46. *Id.* (citation omitted).

47. *Id.* at 1201.

48. *Id.* at 1205 (emphasis in original) (footnote omitted).

49. *Id.*

50. *Id.* at 1202 (quoting *Police Dep’t of Chicago v. Mosley*, 408 U.S. 92, 95 (1972)).

51. *Id.* (quoting *Mosley*, 408 U.S. at 95 (quoting *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 270 (1964))).

52. *Id.* (quoting *Mosley*, 408 U.S. at 95 (quoting *Sullivan*, 376 U.S. at 270)).

53. *Id.* at 1201.

B. Fighting Words

The Court in *Chaplinsky v. New Hampshire* defined “fighting words” as “those which by their very utterance inflict injury or tend to incite an immediate breach of the peace.”⁵⁴ The Court opined that,

[i]t has been well observed 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 that may be derived from them is clearly outweighed by the social interest in order and morality.⁵⁵

Over fifty years after *Chaplinsky*, in *R.A.V. v. St. Paul*, the city of St. Paul issued a Bias-Motivated Crime Ordinance, which read:

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.⁵⁶

The Court held that “[d]isplays containing abusive invective, no matter how vicious or severe, are permissible unless they are addressed to one of the specified disfavored topics.”⁵⁷ It explained that “[t]he First Amendment does not permit [the city] to impose special prohibitions on those speakers who express views on disfavored subjects.”⁵⁸ Thus, the ordinance violated the ban on subject-matter discrimination because it singled out “race, color, creed, religion or gender.”⁵⁹

This case is indicative of the Supreme Court’s aversion to viewpoint discrimination and subject-matter discrimination. Moreover, it is evidence that content-based regulations must pass the strict scrutiny test.⁶⁰ Yet, there are exceptions that permit the government to prohibit speech.⁶¹ First, “[w]hen the basis for the content discrimination consists entirely of the very reason

54. *Chaplinsky v. New Hampshire*, 315 U.S. 568, 572 (1942).

55. *Id.* (footnote omitted).

56. *R.A.V. v. St. Paul*, 505 U.S. 377, 380 (1992) (citation omitted).

57. *Id.* at 391.

58. *Id.* (citations omitted).

59. *Id.*

60. *Reed v. Town of Gilbert*, 135 S. Ct. 2218, 2226–28 (2015).

61. *See infra* notes 62–68 and accompanying text.

the entire class of speech at issue is proscribable, no significant danger of idea or viewpoint discrimination exists.”⁶² The Court explained that “[s]uch a reason, having been adjudged neutral enough to support exclusion of the entire class of speech from First Amendment protection, is also neutral enough to form the basis of distinction within the class.”⁶³ Second, the Court previously dealt with so-called “secondary effects” of speech-prohibitory regulations,⁶⁴ and in *R.A.V.*, explained that when “secondary effects” are implicated, the regulation is “*justified* without reference to the content of the . . . speech.”⁶⁵ The Court echoed its previous sentiment that “[l]isteners’ reactions to speech are not the [intended] ‘secondary effects’” as “[t]he emotive impact of speech on its audience is not a ‘secondary effect.’”⁶⁶ Third, conduct restrictions may be allowed “since words can in some circumstances violate laws directed not against speech but against conduct . . . , a particular content-based subcategory of a proscribable class of speech can be swept up incidentally within the reach of a statute directed at conduct rather than speech.”⁶⁷ Finally, the Court recognizes a catchall exception for speech “so long as the nature of the content discrimination is such that there is no realistic possibility that official suppression of ideas is afoot.”⁶⁸

The form of judicial review known as “strict scrutiny” requires a two-prong test.⁶⁹ First, courts determine if the underlying governmental objectives are “compelling.”⁷⁰ As the government is impeding upon a core constitutional right, “only the most pressing circumstances can justify the government action.”⁷¹ Next, courts analyze “if the law is a narrowly tailored means of furthering those governmental interests.”⁷² In order to be narrowly tailored, the law must “capture within its reach no more activity (or less) than is necessary to advance those compelling ends.”⁷³ That is to say that it “must

62. *R.A.V. v. St. Paul*, 505 U.S. 377, 388 (1992).

63. *Id.*

64. *See Renton v. Playtime Theatres*, 475 U.S. 41, 47 (1986) (distinguishing between concerns about the content of adult movie theaters and “the *secondary effects* of such theaters on the surrounding community” (emphasis in original)).

65. *R.A.V.*, 505 U.S. at 389 (emphasis in original) (quoting *Renton*, 475 U.S. at 48).

66. *Id.* at 394 (quoting *Boos v. Barry*, 485 U.S. 312, 321 (1988)).

67. *Id.* at 389 (citations omitted).

68. *Id.* at 390.

69. Adam Winkler, *Fatal in Theory and Strict in Fact: An Empirical Analysis of Strict Scrutiny in the Federal Courts*, 59 VAND. L. REV. 793, 800 (2006).

70. *Id.*

71. *Id.* (footnote omitted).

72. *Id.*

73. *Id.*

be the ‘least restrictive alternative’ available to pursue those ends.”⁷⁴ By analyzing the “‘fit’ between the ends and the means,” courts can assess the veracity of the government’s alleged goal.⁷⁵

The Supreme Court has explained that “the reason why fighting words are categorically excluded from the protection of the First Amendment is not that their content communicates any particular idea, but that their content embodies a particularly intolerable (and socially unnecessary) *mode* of expressing *whatever* idea the speaker wishes to convey.”⁷⁶ They are “personally abusive epithets which, when addressed to the ordinary citizen, are, as a matter of common knowledge, inherently likely to provoke violent reaction.”⁷⁷

In *Snyder v. Phelps*, the father of a deceased military service member brought an action against a fundamentalist church and its members who held an anti-homosexual demonstration near the service member’s funeral.⁷⁸ The Court explained that speech is protected if it is a matter of public concern, and that “[d]eciding whether speech is of public or private concern require[d] [it] to examine the ‘content, form, and context’ of that speech.”⁷⁹ In addition, the Court felt that liability hinged on the content—not the location—of the speech.⁸⁰ Thus, it stated that the church’s choice of “where and when” to conduct its picketing was “subject to reasonable time, place, or manner restrictions.”⁸¹ The Court “identified a few limited situations where the location of targeted picketing can be regulated under provisions that . . . [are] content neutral.”⁸² As a result, content-based restrictions receive the highest level of scrutiny and content-based speech enjoys the most protection, but content-neutral “‘time, place and manner’ regulations” differ and allow certain speech to be punished.⁸³

74. *Id.* at 800–01 (footnote omitted).

75. *Id.* at 801.

76. *R.A.V. v. St. Paul*, 505 U.S. 377, 393 (1992) (emphasis in original).

77. *Cohen v. California*, 403 U.S. 15, 20 (1971) (citation omitted).

78. *Snyder v. Phelps*, 562 U.S. 443, 448–50 (2011).

79. *Id.* at 453 (internal quotes omitted).

80. *Id.* at 454–56.

81. *Id.* at 456 (citation omitted).

82. *Id.* at 457.

83. *See Police Dep’t of Chicago v. Mosley*, 408 U.S. 92, 98 (1972) (explaining that content-based restrictions must pass strict scrutiny, but content-neutral “time, place and manner” restrictions may be upheld if they are “necessary to further significant governmental interests,” a less demanding level of scrutiny).

C. True Threats

The Court in *Virginia v. Black* defined a “true threat” as a statement “where the speaker means to communicate a serious expression of an intent to commit an act of unlawful violence to a particular individual or group of individuals.”⁸⁴ Thus, the speaker does not actually have to intend to carry out the threat, but instead the prohibition “protects individuals from the fear of violence and the disruption that fear engenders, as well as from the possibility that the threatened violence will occur.”⁸⁵ Here, the harm is not the subsequent violence, but the fear instilled by the threat itself, and this harm happens immediately. One enumerated form of a true threat is intimidation “where a speaker directs a threat to a person or group of persons with the intent of placing the victim in fear of bodily harm or death.”⁸⁶ *Virginia* involved the burning of a cross on someone’s lawn, and the Court cited to the pernicious “history of cross burning in this country” to show “that cross burning is often intimidating, intended to create a pervasive fear in victims that they are a target of violence.”⁸⁷ Cross burning cannot be considered a fighting word because it is not intended or likely to invoke a violent response. Here, the Court held that the government cannot prohibit all cross burning, but the government can prohibit cross burning done with the intent to threaten or intimidate, as this would qualify as a “true threat.”⁸⁸ But there must be proof that the cross burning was a true threat.⁸⁹

The Supreme Court jurisprudence surrounding the First Amendment has evolved in the past century.⁹⁰ The Court has carved out exceptions to free speech when that speech is an incitement to imminent lawless action, a fighting word, or a true threat.⁹¹ The true threat doctrine parts ways from the incitement doctrine, as it does not mandate an imminency requirement.⁹² And one of the purposes in banning fighting words is to protect the speaker from the natural harm that could arise as a result of uttering the fighting words.⁹³ In its collective jurisprudence on the subject, the Supreme Court has yet to

84. *Virginia v. Black*, 538 U.S. 343, 359 (2003).

85. *Id.* at 359–60 (citation omitted).

86. *Id.* at 344.

87. *Id.* at 360.

88. *Id.* at 359–60.

89. *Id.* at 368 (striking down the portion of the Virginia statute providing that cross burning served as prima facie evidence of an intent to intimidate).

90. *See supra* notes 14–89 and accompanying text.

91. *See supra* notes 23–89 and accompanying text.

92. *See supra* notes 84–89 and accompanying text.

93. *See generally* *Cohen v. California*, 403 U.S. 15, 20 (1971) (implying harmful words could invoke harmful action).

fashion a cohesive model for how to constitutionally proscribe hate speech to ensure that free speech protections no longer lead to individuals feeling harassed or threatened, or as though they may be the victims of violence.

II. HATE SPEECH ONLINE AND HATE CRIMES IN PERSON

As of 2019, approximately 70% of Americans use at least one social media website.⁹⁴ Social media has been described as “web-based communication tools that enable people to interact with each other by [both] sharing and consuming information.”⁹⁵ Features common to most social media platforms include personal user accounts, profile pages, the ability to make original posts, friends/followers, groups, searchable hashtags, Like buttons, comment sections, news feeds, personalization, and notifications.⁹⁶

What commences as hate speech on social media often leads to hate crimes in person.⁹⁷ A 2018 study found that over half of Americans surveyed said that they have been personally harassed online, with more than one-third reporting severe harassment.⁹⁸ Sixty-two percent of Americans surveyed viewed online harassment “as a major problem” and seventy-nine percent feel that online services “have a duty to step in when harassment occurs on their platforms.”⁹⁹ The Simon Wiesenthal Center released its 2019 Digital Terrorism & Hate Report Card, which provided a letter grade to social media platforms regarding their efforts to combat hate, with the highest grade given a B+ and the lowest, a barely passing D+.¹⁰⁰ The Center estimates that there are “approximately 15,000 problematic Web sites, social networking pages, forums and newer online technology games and apps” that are used to transmit acrimonious rhetoric.¹⁰¹ Moreover, there are currently approximately 1,020 hate groups operating in some fashion across the

94. *Social Media Fact Sheet*, *supra* note 5.

95. Daniel Nations, *What Is Social Media?*, LIFEWIRE, <https://www.lifewire.com/what-is-social-media-explaining-the-big-trend-3486616> (last updated Nov. 24, 2019) (emphasis omitted).

96. *Id.*

97. *See generally* Evanna Hu & Hugh Brooks, *When Radical Online Political Speech Transitions to Physical Violence*, PAC. STANDARD (Aug. 24, 2018), <https://psmag.com/social-justice/when-radical-speech-gets-physical> (displaying the vulnerability of online platforms and how the opposition may retaliate in the real world).

98. *Online Hate and Harassment*, *supra* note 6.

99. Duggan, *supra* note 8.

100. *Digital Terrorism & Hate Report Card*, SIMON WIESENTHAL CTR., <http://digitalhate.net/inicio.php> (last visited Dec. 11, 2019).

101. *Social Media Must Do More*, *supra* note 7.

country.¹⁰² And “[f]or the first time in . . . years, hate groups were [recently] found in all 50 states.”¹⁰³

In a 2017 Supreme Court case, Justice Anthony Kennedy said that “we cannot appreciate yet [cyberspace’s] full dimensions and vast potential to alter how we think, express ourselves, and define who we want to be.”¹⁰⁴ Mainstream social networking sites struggle with how to handle hate groups on their platforms, while balancing “the right to share and debate ideas with the responsibility to protect society against potential attacks.”¹⁰⁵ The Unite the Right Rally in Charlottesville in August 2017 was a chief illustration of how the ease of online speech can vivify hate groups in a way that leads to violence.¹⁰⁶ These groups used the Internet to recruit and circulate propaganda.¹⁰⁷ They used content-specific websites and chat rooms, as well as mainstream social media platforms, to plan the rally.¹⁰⁸ Over the course of the two months before the rally, “there were over 35,000 vitriolic messages on 44 ‘channels’ on the Charlottesville 2.0 server on the gaming platform Discord.”¹⁰⁹ After the rally, these groups attempted to use platforms to raise money for their legal defenses.¹¹⁰ Some platforms deactivated accounts, but users responded accordingly by finding new ways to spread their hate-fueled messages.¹¹¹

Facebook CEO Mark Zuckerberg said that Facebook has always “taken down any post that promotes or celebrates hate crimes or acts of terrorism” and that it would continue to “take down threats of physical harm.”¹¹² But activist organizations contend that these measures are still not sufficient.¹¹³ In 2016, Facebook was provided with a list of over 200 hate groups that were on the site.¹¹⁴ The list included pages devoted to denying the Holocaust, as well as white-nationalist, anti-Muslim, black-separatist, and Neo-Nazi

102. *Hate Map*, S. POVERTY LAW CTR., <https://www.splcenter.org/hate-map> (last visited Dec. 12, 2019).

103. *Five Ways Hate Speech Spreads Online*, CNN (May 2, 2018), <https://www.cnn.com/2018/05/02/us/hate-groups-using-internet/index.html>.

104. *Packingham v. North Carolina*, 137 S. Ct. 1730, 1736 (2017).

105. *Five Ways Hate Speech Spreads Online*, *supra* note 103.

106. Hu & Brooks, *supra* note 97.

107. *Id.*

108. *Id.*

109. *Id.*

110. *Id.*

111. *Id.*

112. Heather Kelly, *Hate Groups on Facebook: Why Some Get to Stay*, CNN BUSINESS (Aug. 17, 2017), <https://money.cnn.com/2017/08/17/technology/culture/facebook-hate-groups/index.html>.

113. *Id.*

114. *Id.*

groups.¹¹⁵ As of 2017, Facebook had removed only 57 of the recommended groups.¹¹⁶ The website has its own internal guidelines about what constitutes a hate group, and “[a] person or group must threaten violence, declare it has a violent mission or actually take part in acts of violence.”¹¹⁷ Facebook deletes approximately 66,000 hateful posts per week.¹¹⁸

In March 2019, a massacre at two mosques in Christchurch, New Zealand left 50 dead and dozens more injured at the hands of a white supremacist who live-streamed the attack on Facebook.¹¹⁹ Less than two weeks later, Facebook announced a ban on “praise, support and representation of white nationalism and . . . separatism on Facebook and Instagram.”¹²⁰ Facebook wrote that it is “clear that these concepts are deeply linked to organized hate groups and have no place on our services.”¹²¹ It also conceded the “need to get better and faster at finding and removing hate from our platforms.”¹²² Additionally, the company stated that it would “start connecting people who search for terms associated with white supremacy to resources focused on helping people leave behind hate groups.”¹²³

Furthermore, Facebook has decided to establish an external, independent oversight board with the power to hear appeals to evaluate and override the company’s decisions on taking down hateful speech.¹²⁴ This leaves Facebook to construct a judicial system for a constituency of 2.3 billion, a feat that includes identifying who will decide what cases can be appealed and whether the decisions will be published like Supreme Court decisions.¹²⁵ The initial court would consist of forty paid judges—neither affiliated with Facebook nor the government—“serving on a part-time basis for three-year terms.”¹²⁶ The court would be equipped with the power “to overrule Facebook’s content moderators, but not to rewrite the company’s Community Standards.”¹²⁷ The

115. *Id.*

116. *Id.*

117. *Id.*

118. *Id.*

119. *Australian Arrested Over New Zealand Shooting Massacre to Face 50 Murder Charges—Police*, REUTERS (Apr. 3, 2019), <https://uk.reuters.com/article/uk-newzealand-shooting-idUKKCN1RG09E>.

120. *Standing Against Hate*, FACEBOOK NEWSROOM (Mar. 27, 2019), <https://newsroom.fb.com/news/2019/03/standing-against-hate/>.

121. *Id.*

122. *Id.*

123. *Id.*

124. Simon Van Zuylen-Wood, “*Men Are Scum: Inside Facebook’s War on Hate Speech*,” VANITY FAIR (Feb. 26, 2019), <https://www.vanityfair.com/news/2019/02/men-are-scum-inside-facebook-war-on-hate-speech>.

125. *Id.*

126. *Id.*

127. *Id.*

judges would also be provided with context when deciding cases in order to best understand the motivations behind each post.¹²⁸

Another major social media platform, Twitter, also reacted following the events in Charlottesville.¹²⁹ Twitter CEO Jack Dorsey described more aggressive policies, including treating hateful imagery and hate symbols on Twitter as “sensitive media,” meaning that “the content will be blurred and users will need to manually opt in to view” it.¹³⁰ But Twitter did not explicate on what it considers a “hate symbol.”¹³¹ Twitter’s policy on hateful conduct states that, “[y]ou may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease.”¹³² It also does not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories.¹³³ It also does not permit “hateful images or symbols” in profile images or profile headers, or using a “username, display name, or profile bio to engage in abusive behavior, such as targeted harassment or expressing hate towards a person, group, or protected category.”¹³⁴ The social networking platform says that it is “committed to combating abuse motivated by hatred, prejudice or intolerance, particularly abuse that seeks to silence the voices of those who have been historically marginalized.”¹³⁵ Twitter’s policy applies when users make “violent threats”; “[w]ish[], hop[e] or call[] for serious harm on a person or group of people”; make “[r]eferences to mass murder, violent events, or specific means of violence where protected groups have been the primary targets or victims”; “[i]ncit[e] fear about a protected category”; “[r]epeat[] . . . slurs, epithets, racist and sexist tropes, or other content that degrades someone”; or share “[h]ateful imagery.”¹³⁶ When determining the penalty for violating this policy, Twitter considers “a number of factors,” such as “the severity of the violation and an individual’s previous record of rule violations.”¹³⁷ It may ask the user to remove the content in question and “serve a period of time in read-only mode before they can Tweet again.”¹³⁸

128. *Id.*

129. *Five Ways Hate Speech Spreads Online*, *supra* note 103.

130. *Id.*

131. *Id.*

132. *Hateful Conduct Policy*, TWITTER, <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (last visited Dec. 12, 2019).

133. *Id.*

134. *Id.*

135. *Id.*

136. *Id.* (emphasis omitted).

137. *Id.*

138. *Id.*

Those who have subsequent violations will be subjected to longer read-only periods and may eventually find their accounts permanently suspended.¹³⁹ If an account “is engaging primarily in abusive behavior, or is deemed to have shared a violent threat,” it will be permanently suspended.¹⁴⁰

YouTube, and its owner Google, also vowed to do more to remove hateful content.¹⁴¹ Currently, YouTube states that it removes “content promoting violence or hatred against individuals or groups” based on the following attributes: age, disability, ethnicity, gender, nationality, race, immigration status, religion, sex, sexual orientation, or veteran status.¹⁴² The video sharing platform instructs users not to post content if their purpose is to “[e]ncourage violence against individuals or groups” or “incite hatred” based on the aforementioned attributes, or to dehumanize individuals or groups by calling them subhuman, comparing them to animals, insects, pests, disease, or any other non-human entity.¹⁴³ Channels are terminated after receiving three strikes.¹⁴⁴ But YouTube promised to take a “tougher stance” on videos that contain, for instance, “inflammatory religious or supremacist content” by including an “interstitial warning” and not allowing the offensive videos to be “monetized, recommended or eligible for comments or user endorsements.”¹⁴⁵

In addition to posting content online, hate groups also utilize the Internet to raise funds ahead of any potential real-world events.¹⁴⁶ PayPal attempts to ensure that its services are not used to accept payments or donations that “promote hate, violence or racial intolerance.”¹⁴⁷ Crowd-funding platform GoFundMe has also shut down campaigns to raise money for James Fields, the man convicted of first-degree murder after he drove into a crowd of people at a rally in response to Charlottesville, amongst other fundraisers supporting white-nationalist or neo-Nazi groups.¹⁴⁸ Yet other groups

139. *Id.*

140. *Id.*

141. *Five Ways Hate Speech Spreads Online*, *supra* note 103.

142. *Hate Speech Policy*, YOUTUBE HELP, <https://support.google.com/youtube/answer/2801939?hl=en> (last visited Dec. 11, 2019) (“If you make a video that compares people to animals or bugs because of their race or nationality, that could result in a strike on your channel.”).

143. *Id.*

144. *Id.*

145. *Five Ways Hate Speech Spreads Online*, *supra* note 103.

146. Kaya Yurieff, *PayPal Is Quietly Cracking Down on White-Supremacist Accounts*, CNN BUSINESS (Aug. 16, 2017), <https://money.cnn.com/2017/08/16/technology/paypal-alt-right-accounts/index.html>.

147. *Id.*

148. *Id.*

circumvent these hurdles by using alt-right-focused platforms or cryptocurrency.¹⁴⁹

Furthermore, many hate groups have their own websites, hosted via companies like GoDaddy and Google Domains.¹⁵⁰ Unfortunately though, web hosts do not have any legal liability or responsibility unless sites are violating federal crimes or infringing on another's intellectual property rights.¹⁵¹

Finally, the so-called "dark web," is a hidden network that can only be accessed through special browsers, not traditional search engines.¹⁵² There are pros and cons for hate groups that are forced onto the dark web. First, there is overall less digital policing, meaning that they have more freedom and more lax standards regarding what can be posted and what will and will not be removed.¹⁵³ On the other hand, the audiences are typically much smaller than what mainstream social media accounts will receive or what Internet users can find using a traditional search engine.¹⁵⁴

Forty civil- and human-rights organizations have joined forces to inform Internet companies that they want to "CHANGE THE TERMS."¹⁵⁵ They view websites as "new tools to those who want to hatefully threaten, harass, intimidate, defame, or even violently attack people different from themselves."¹⁵⁶ As such, this permits organizers "to mobilize, organize, fundraise, and normalize racism, sexism, bigotry, and xenophobia."¹⁵⁷ The Change the Terms policies were drafted by the Center for American Progress, Color of Change, Free Press, the Lawyers' Committee for Civil Rights Under Law, the National Hispanic Media Coalition, and the Southern Poverty Law Center.¹⁵⁸ They hope to formulate a "set of uniform model policies that civil and human rights organizations could point to as best practices" and "that these new policies can set a benchmark to measure the progress of major tech companies, as well as a guide for newer companies wrestling with some of these issues for the first time."¹⁵⁹ A company who adopts these policies

149. *Five Ways Hate Speech Spreads Online*, *supra* note 103.

150. Yurieff, *supra* note 146.

151. *Five Ways Hate Speech Spreads Online*, *supra* note 103.

152. *Id.*

153. *Id.*

154. *Id.*

155. *FAQS*, CHANGE THE TERMS.ORG, <https://www.changethetterms.org/faqs> (last visited Dec. 11, 2019).

156. *Id.*

157. *Id.*

158. *Id.*

159. *Id.*

“commits to not allowing [its] services to be used for hateful activities.”¹⁶⁰ And the policies include provisions for enforcement, the right to appeal, transparency, evaluation and training, governance and authority, and state actors, bots, and trolls.¹⁶¹

While hate speech is still constitutionally protected in many respects, the federal government has enacted laws that criminalize *acts* of hatred.¹⁶² A hate crime is “a traditional offense like murder, arson, or vandalism with an added element of bias.”¹⁶³ “[T]he FBI has defined a hate crime as a ‘criminal offense against a person or property motivated in whole or in part by an offender’s bias against a race, religion, disability, sexual orientation, ethnicity, gender, or gender identity.’”¹⁶⁴

In 1968, President Lyndon Johnson signed the first federal hate-crime statute into law, which

made it a crime to use, or threaten to use, force to willfully interfere with any person because of race, color, religion, or national origin, and because the person is participating in a federally protected activity, such as public education, employment, jury service, travel, or the enjoyment of public accommodations, or helping another person to do so.¹⁶⁵

The same year,¹⁶⁶ the Fair Housing Act included a provision that made it a crime to use force or threaten to use force to interfere with housing rights on the basis of race, color, religion, sex, or national origin.¹⁶⁷ Nearly 30 years later, Congress passed the Church Arson Prevention Act, whereby it is a “crime to deface, damage, or destroy religious real property, or interfere with a person’s religious practice, in situations affecting interstate commerce.”¹⁶⁸ In addition, the Act “bars defacing, damaging, or destroying religious

160. *Id.* (citation omitted) (defining “hateful activities” as actions “that incite or engage in violence, intimidation, harassment, threats, or defamation targeting an individual or group based on their actual or perceived race, color, religion, national origin, ethnicity, immigration status, gender, gender identity, sexual orientation, or disability”).

161. *Id.* (discussing each provision in detail).

162. *See infra* notes 163–75 and accompanying text.

163. *See Hate Crimes, supra* note 9.

164. *Id.*

165. *Hate Crime Laws*, U.S. DEP’T OF JUSTICE (last updated Mar. 7, 2019), <https://www.justice.gov/crt/hate-crime-laws>; Violent Interference with Federally Protected Rights; 18 U.S.C. § 245 (2012).

166. *Hate Crime Laws, supra* note 165; Fair Housing Act, 42 U.S.C. § 3631 (1996).

167. Paula A. Franzese & Stephanie J. Beach, *Promises Still to Keep: The Fair Housing Act Fifty Year Later*, 40 CARDOZO L. REV. 1207, 1223 n.103 (2019).

168. *Hate Crime Laws, supra* note 165; Church Arson Prevention Act, 18 U.S.C. § 247 (1994).

property because of the race, color, or ethnicity of persons associated with the property.”¹⁶⁹

In 2009, President Barack Obama signed the Matthew Shepard and James Byrd Jr. Hate Crimes Prevention Act into law.¹⁷⁰ This law expands the federal definition of hate crimes and provides funding and technical assistance to state, local, and tribal jurisdictions to assist them in more effectively investigating and prosecuting hate crimes.¹⁷¹ The Act criminalizes willfully causing bodily injury (or attempting to do so with fire, firearm, or other dangerous weapon) when: (1) the crime was committed because of the “actual or perceived race, color, religion, or national origin” of any person; or (2) the crime was committed because of the “actual or perceived religion, national origin, gender, sexual orientation, gender identity, or disability” of any person and the crime affected interstate or foreign commerce or occurred within federal special maritime and territorial jurisdiction.¹⁷² This law “removed then existing jurisdictional obstacles to prosecutions of certain race- and religion-motivated violence, and added new federal protections against crimes based on gender, disability, gender identity, or sexual orientation.”¹⁷³ Now, before the Civil Rights Division prosecutes a hate crime, the Attorney General must certify in writing that:

(A) the State does not have jurisdiction; (B) the State has requested that the Federal Government assume jurisdiction; (C) the verdict or sentence obtained pursuant to State charges left demonstratively unvindicated the Federal interest in eradicating bias-motivated violence; or (D) a prosecution by the United States is in the public interest and necessary to secure substantial justice.¹⁷⁴

It is also a crime to conspire “to injure, oppress, threaten, or intimidate any person . . . in the free exercise or enjoyment of any right or privilege secured to him by the Constitution or laws of the United States, or because of his having so exercised the same.”¹⁷⁵

Moreover, in *Wisconsin v. Mitchell*, the Supreme Court held that enhanced sentencing for bias-motivated crimes does not impinge on a

169. *Hate Crime Laws*, *supra* note 165.

170. *Id.*; The Matthew Shepard and James Byrd Jr. Hate Crimes Prevention Act of 2009, 18 U.S.C. § 249 (2012).

171. 18 U.S.C. § 249.

172. *Id.*

173. *Hate Crime Laws*, *supra* note 165.

174. 18 U.S.C. § 249.

175. Conspiracy Against Rights, 18 U.S.C. § 241 (2012).

defendant's First Amendment rights.¹⁷⁶ Thus, government-neutrality for topics or viewpoints only applies to speech, not to conduct. "The First Amendment . . . does not prohibit the evidentiary use of speech to establish the elements of a crime or to prove motive or intent. Evidence of a defendant's previous declarations or statements is commonly admitted in criminal trials subject to evidentiary rules dealing with relevancy, reliability, and the like."¹⁷⁷ This case established groundbreaking precedent concerning the crosshairs between free-speech and hate-crime legislation.¹⁷⁸

The key is that these laws criminalize hate *crimes*, but not hate *speech*. Consequently, speech must lead to actions before it can be punished. This is compounded by the fact that these statutes are not focused on the Internet. However, there are other federal laws that regulate the online space.¹⁷⁹ Congress enacted the Communications Decency Act of 1996 (CDA) to fight indecency and obscenity online.¹⁸⁰ Section 230 of the Act states that, "[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."¹⁸¹ That is, online intermediaries that host or republish speech are immune from liability for statements made by others. These intermediaries include not only regular Internet Service Providers (ISPs), but also a range of "interactive computer service[]" providers, which are online services that publish third-party content.¹⁸² The Act also protects bloggers from comments posted by readers.¹⁸³

In 1998, Congress enacted the Children's Online Privacy Protection Act (COPPA) in order to protect the online privacy of children under the age of thirteen.¹⁸⁴ It "imposes certain requirements on operators of websites or online services directed to children under 13 years of age, and on operators of other websites or online services that have actual knowledge that they are collecting personal information online from a child under 13 years of age."¹⁸⁵

176. *Wisconsin v. Mitchell*, 508 U.S. 476, 490 (1993).

177. *Id.* at 489.

178. *Id.*

179. *See infra* notes 180–85 and accompanying text.

180. Communications Decency Act of 1996, 47 U.S.C. § 230 (2012).

181. *Id.*

182. *Id.*

183. *Section 230 of the Communications Decency Act*, ELEC. FRONTIER FOUND., <https://www.eff.org/issues/cda230> (last visited Dec. 14, 2019) (discussing favorably wide-ranging real-world impacts of section 230).

184. Children's Online Privacy Protection Act, 15 U.S.C. §§ 6501–06 (2012).

185. *Children's Online Privacy Protection Rule ("COPPA")*, FED. TRADE COMM'N, <https://www.ftc.gov/enforcement/rules/rulemaking-regulatory-reform-proceedings/childrens-online-privacy-protection-rule> (last visited Dec. 14, 2019).

The question lingers as to why this legislation, which is focused on protecting children's privacy, does not aim to protect children from such speech.

In seeking alternative resources, the intentional infliction of emotional distress is a common law tort that permits recovery for severe emotional distress caused by "extreme and outrageous conduct" or conduct that exceeds "all bounds . . . of decency."¹⁸⁶ If the conduct in question is continuous or repetitive in nature, the likelihood that it is outrageous increases.¹⁸⁷ In addition, if the speaker is aware of a particular emotional weakness or sensitivity on the part of his or her target and exploits that, this person's action also qualifies as outrageous.¹⁸⁸ Another tort is the negligent infliction of emotional distress, which requires a potential defendant to take precautions to avoid disturbing another's mental or emotional tranquility, either intentionally or unintentionally.¹⁸⁹ The standard of care in a negligence action is that which a reasonable person should exercise under the circumstances.¹⁹⁰ Here, though, physical manifestations of the stress should be shown.¹⁹¹ The difficulty that lies in attempting to hold platforms legally responsible for not removing hate under a common law theory is that secondary or vicarious liability applies when there is an agency relationship, and the "right and ability to supervise" is done by the "beneficiary of that exploitation."¹⁹² The business model of these sites is that users create the

186. RESTATEMENT (SECOND) OF TORTS: OUTRAGEOUS CONDUCT CAUSING SEVERE EMOTIONAL DISTRESS § 46 (AM. LAW INST. 1965) ("One who by extreme and outrageous conduct intentionally or recklessly causes severe emotional distress to another is subject to liability for such emotional distress, and if bodily harm to the other results from it, for such bodily harm."); *id.* at cmt. d.

187. "Liability has been found only where the conduct has been so outrageous in character, and so extreme in degree, as to go beyond all possible bounds of decency, and to be regarded as atrocious, and utterly intolerable in a civilized community. Generally, the case is one in which the recitation of the facts to an average member of the community would arouse his resentment against the actor, and lead him to exclaim, 'Outrageous!'" *Id.* at cmt. d.

188. "The extreme and outrageous character of the conduct may arise from the actor's knowledge that the other is peculiarly susceptible to emotional distress, by reason of some physical or mental condition or peculiarity. The conduct may become heartless, flagrant, and outrageous when the actor proceeds in the face of such knowledge, where it would not be so if he did not know. It must be emphasized again, however, that major outrage is essential to the tort; and the mere fact that the actor knows that the other will regard the conduct as insulting, or will have his feelings hurt, is not enough." *Id.* at cmt. f.

189. RESTATEMENT (SECOND) OF TORTS § 312-13 (AM. LAW INST. 1965).

190. *Id.*

191. *Id.*

192. *Shapiro, Bernstein & Co. v. H.L. Green Co.*, 316 F.2d 304, 307 (2d Cir. 1963) (citations omitted) ("When the right and ability to supervise coalesce with an obvious and direct financial interest[,] . . . even in the absence of actual knowledge that the copyright monopoly is being impaired[,] . . . the purposes of copyright law may be best effectuated by the imposition of liability upon the beneficiary of that exploitation.").

content, not the websites.¹⁹³ But one can still argue that social media platforms “supervise” the content—especially as there exists the ability to report posts that do not abide by stipulated policies—and simultaneously “benefit” from having more users take advantage of their services.¹⁹⁴

The current climate is indicative of the dangers of hate speech and the very real and very frightening crimes that can happen because of the propagation of that speech. Therefore, it is due time for legal and regulatory reform, both on the part of private and governmental action.

III. A PRESCRIPTION FOR LEGAL AND REGULATORY REFORM

One of the seminal United States Supreme Court defamation cases explained that, “[u]nder the First Amendment there is no such thing as a false idea. However pernicious an opinion may seem, we depend for its correction not on the conscience of judges and juries but on the competition of other ideas.”¹⁹⁵ There are those who claim that hate speech promotes dialogue and debate and is not a false statement of fact like libel.¹⁹⁶ Yet, these claims ignore the crucial fact that hate speech can, and often does, lead to hate crimes of physical acts of violence targeting specific individuals or communities.¹⁹⁷ According to the FBI, in 2018, there were 7,036 single-bias incidents involving 8,646 victims.¹⁹⁸ The largest bias-motivation was on the basis of race/ethnicity/ancestry, followed by religion, then sexual orientation, gender identity, disability, and gender.¹⁹⁹

Hate speech can provide “vocabulary and grammar depicting a common enemy . . . [and establish] a mutual interest in trying to rid society of the

193. See Ben Gilbert, *How Facebook Makes Money from Your Data*, in *Mark Zuckerberg's Words*, BUS. INSIDER (Apr. 11, 2018), <https://www.businessinsider.com/how-facebook-makes-money-according-to-mark-zuckerberg-2018-4> (explaining how Facebook utilizes users' content to sell advertisements).

194. See *Shapiro, Bernstein & Co.*, 316 F.2d at 307.

195. *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 339–40 (1974) (footnote omitted).

196. See *Anti-Defamation League v. Fed. Commc'ns Comm.*, 403 F.2d 169, 174 (Skelly Wright, J., concurring) (“Under the law of libel, defamation of a broad group or class is not usually actionable. And this kind of speech, detestable as some of its anti-Semitic and racist aspects may be, approaches the area of political and social commentary.” (footnote omitted)).

197. Hatzipanagos, *supra* note 9 (“The digital world gives white supremacists a safe space to explore extreme ideologies and intensify their hate without consequence Their rage can grow under the radar until the moment when it explodes in the real world.”).

198. *2018 Hate Crime Statistics*, U.S. DEP'T OF JUSTICE, <https://www.justice.gov/hatecrimes/hate-crime-statistics> (last visited Dec. 14, 2019).

199. *Id.*

designated pest.”²⁰⁰ An essential inquiry, as a starting point, is determining what the harm is in hate speech. First, hate speech allows for radicalization and for those with hateful views to feel normal.²⁰¹ When normalization occurs, ideas and viewpoints can move from the fringe to the conventional.²⁰² By expanding due to the infinite vastness of the online world, these beliefs can gain traction and momentum.²⁰³ This allows an audience to grow, consisting of those who identify with these feelings and emotions and the expression behind the speech.²⁰⁴ What is troublesome is that this behavior is abetted by algorithms on platforms that spiral users down a so-called “feedback loop,” showcasing more and more deleterious content.²⁰⁵ Once down the rabbit hole of toxic propaganda, it is hard to climb back up.

It is not difficult to draw a parallel between this phenomenon and the radicalization of extremists and terrorists.²⁰⁶ But, while this is akin to the spreading of terroristic ideologies, the online world has taken a harsher stance against a perception of inadvertently promoting terrorism-sympathizing material.²⁰⁷ This is compounded by the fact that federal agencies do not even supply common definitions for “domestic terrorism” or a “domestic terrorist.”²⁰⁸ And these governmentally-circulated lists provide social media companies with a starting point on how to police their community content.²⁰⁹ Many companies also rely on a shared database of terrorist content, coordinated through the Global Internet Forum to Counter Terrorism, to expedite the removal of toxic posts.²¹⁰ Crucially, the Forum is capable of

200. ALEXANDER TESIS, *DESTRUCTIVE MESSAGES: HOW HATE SPEECH PAVES THE WAY FOR HARMFUL SOCIAL MOVEMENTS* 102 (2002).

201. See Hatzipanagos, *supra* note 9 (“The digital world gives white supremacists a safe space to explore ideologies and intensify their hate without consequence . . .”).

202. See *id.* (explaining how “hateful content [on 4chan] spread like a virus to more mainstream sites such as Facebook, Twitter and Instagram through shared memes and retweets, where they reach much larger audiences”).

203. See *id.* (“[E]xtremist groups are able to quickly normalize their messages by delivering a never-ending stream of hateful propaganda to the masses.”).

204. See *id.* (“[Social media] allows people to see other hateful words, slurs and ideas, and those things become normal . . . Norms are powerful because they influence people’s behaviors. If you see a string of slurs, that makes you feel like things are more acceptable.”).

205. The Editorial Board, *The New Radicalization of the Internet*, N.Y. TIMES (Nov. 24, 2018), <https://www.nytimes.com/2018/11/24/opinion/sunday/facebook-twitter-terrorism-extremism.html>.

206. *Id.*

207. See generally, U.N. OFFICE ON DRUGS & CRIME, *THE USE OF THE INTERNET FOR TERRORIST PURPOSES* 5, 6–7 (2012), https://www.unodc.org/documents/frontpage/Use_of_Internet_for_Terrorist_Purposes.pdf (reporting on dangers of terrorist internet use, particularly with respect to recruitment and ideology dissemination).

208. The Editorial Board, *supra* note 205.

209. *Id.*

210. *Id.*

identifying information based upon the information curated by official organizations.²¹¹

Under federal law, the term “domestic terrorism” denotes activities that: (1) “involve acts dangerous to human life that are a violation of the criminal laws of the United States or of any State;” (2) “appear to be intended . . . to intimidate or coerce a civilian population; . . . influence the policy of a government by intimidation or coercion; or . . . to affect the conduct of a government by mass destruction, assassination, or kidnapping;” and (3) “occur primarily within the territorial jurisdiction of the United States.”²¹² Several states have enacted statutes imposing criminal liability for making “terroristic threats.”²¹³ Terrorists do not become indoctrinated from videos instructing people how to make bombs, but rather from speeches by extremists.²¹⁴ But the FBI does not monitor these websites due to restrictions imposed by the Attorney General Guidelines.²¹⁵ These Guidelines were originally devised to prohibit the FBI from investigating speech activities when there is not a call for violence, or from monitoring certain political views.²¹⁶ Instead, only criminal activity could prompt an FBI investigation.²¹⁷ Over the years, the Guidelines have been revised and molded into their current form.²¹⁸ Now, FBI agents are prohibited from proactively monitoring websites unless they suspect criminal activity.²¹⁹ But they can still use public databases and websites to search for evidence of criminality or potential threats to national security.²²⁰

Next, social influence plays a role in weaponizing inflammatory discourse. Social influence is the phenomenon that occurs when someone’s emotions, opinions, or behaviors are either consciously or unconsciously impacted by others.²²¹ Particular communications can cause changes in

211. *Id.*

212. 18 U.S.C. § 2331(5) (2012).

213. See John P. Ludington, Annotation, *Validity and Construction of Terroristic Threat Statutes*, 45 A.L.R. 4th 949 (1986) (delineating state terroristic threat statutes).

214. Nikita Malik, *Hate Speech Still Spreads Too Easily Online: Here’s How to Stop It*, FORBES (Oct. 12, 2018), <https://www.forbes.com/sites/nikitamalik/2018/10/12/hate-speech-still-spreads-too-easily-online-heres-how-to-stop-it/#288309e53821>.

215. David Morgan, *Policing Social Media to Combat Online Radicalization*, CBS NEWS (Oct. 29, 2018), <https://www.cbsnews.com/news/pittsburgh-shooting-gab-policing-social-media-to-combat-online-radicalization/>.

216. *Id.*

217. *Id.*

218. *Id.*

219. *Id.*

220. *Id.*

221. *Personal MBTI Type Analysis*, QUALITIES OF A LEADER (Dec. 26, 2011), <https://web.archive.org/web/20120322041448/http://qualities-of-a-leader.com/personal-mbti-type-analysis>.

attitudes and beliefs that can either be surface level or “integrated into the person’s value system.”²²² Psychologist Herbert Kelman identified three varieties of social influence.²²³ First, compliance is the theory that “an individual accepts influence because he hopes to achieve a favorable reaction from another person or group.”²²⁴ The “satisfaction derived from compliance is due to the *social effect* of accepting influence.”²²⁵ Second, identification occurs “when an individual accepts influence because he wants to establish or maintain a satisfying self-defining relationship to another person or a group.”²²⁶ Now the “satisfaction” is “due to the *act* of conforming as such.”²²⁷ Third, internalization happens “when an individual accepts influence because the content of the induced behavior—the ideas and actions of which it is composed—is intrinsically rewarding.”²²⁸ This time, “satisfaction” is “due to the *content* of the new behavior.”²²⁹

The damage caused because of the power of social influence lies in the affirmation that perpetrators receive when their uploaded content receives likes, shares, views, and comments.²³⁰ Because of the dopamine rush that they will experience, users will feel the increased desire to continue posting such content with the aspiration that others will continue feeding off of this rhetoric and then share it.²³¹

Next, as the “true threats” doctrine explains,²³² people feel threatened or harassed by speech that targets them specifically or the groups they are a part of. The Supreme Court has left much to be desired in its case law on this issue. For starters, questions linger as to whether the speaker must purport to instill fear into the listener, whether the speaker must know that his or her words will instill fear, or, on the other hand, if the heightened risk is sufficient.²³³ Additionally, threats can run the gamut of being specific,

222. Herbert C. Kelman, *Compliance, Identification, and Internalization: Three Processes of Attitude Change*, 2 J. CONFLICT. RES. 51, 51 (1958).

223. *Id.* at 51, 53.

224. *Id.* at 53.

225. *Id.* (emphasis in original).

226. *Id.*

227. *Id.* (emphasis in original).

228. *Id.*

229. *Id.* (emphasis in original).

230. Hal McDonald, *Social Media Affirmation Goes Both Ways*, PSYCHOLOGY TODAY (Sept. 28, 2018), <https://www.psychologytoday.com/us/blog/time-travelling-apollo/201809/social-media-affirmation-goes-both-ways>.

231. Trevor Haynes, *Dopamine, Smartphones & You: A Battle for Your Time*, HARV. U: SCI. IN THE NEWS (May 1, 2018), <http://sitn.hms.harvard.edu/flash/2018/dopamine-smartphones-battle-time/>.

232. *See supra* Part I.C (defining and further discussing the true threats doctrine).

233. Jacob Hillesheim, *When Does Online Harassment Turn into a True Threat?*, REWIRE (July 30, 2018), <https://www.rewire.org/our-future/online-harassment-true-threat/>.

conditional, or general.²³⁴ But it is essential to determine whether, as a society, we want to permit such behavior. The line between threatening speech and hate speech can be quite fine. When a social media user posts an image of a noose next to an African American or a swastika next to a Jewish person, is this a “true threat,”²³⁵ or is it protected rhetoric? And protected by whom—the government or the host site? Any non-governmental entity has the right to restrict speech as it sees fit or to penalize users who make comments that they deem unacceptable.²³⁶ Accordingly, by using the “true threats” doctrine as a model, these platforms should enforce a set of standards where this speech is deemed unprotected and impermissible.

As the internet is a global network, any idea has the potential of reaching far across the globe. “If a hate message influences persons to attack someone, it does not matter whether the victim lives near or far from the location of transmission; the danger is just as great because the Internet is a global network.”²³⁷ As such, it is imperative that actions are taken to restrict the spread of such malignancies. There are both private and governmental options that can be implemented, and each carries a set of positive and negative consequences and its own hurdles to execution and enforcement.²³⁸

A. Private Action

The First Amendment does not apply to the policies of a private company; it only applies to actions taken by federal, state, or local government.²³⁹ Thus, the Supreme Court’s First Amendment jurisprudence does not apply to regulations disseminated by private corporations.²⁴⁰ So, private companies have the ability to make guidelines regarding what is hateful and impermissible, and it is not unconstitutional for them to ban users based upon the content that they are spreading.²⁴¹ In order to have uniformity, all social media sharing sites should adopt these mechanisms. This would not

234. *See id.* (discussing how courts look for expressions of intent to carry out the threat in determining whether a threat is protected speech).

235. *See supra* Part I.C.

236. *See supra* Part II (discussing online hate speech regulation).

237. Alexander Tsesis, *Hate in Cyberspace: Regulating Hate Speech on the Internet*, 38 SAN DIEGO L. REV. 817, 837 (2001).

238. *See infra* Parts III.A, III.B.

239. *See* Stephen K. Wirth, *State Action, Government Speech, and the Narrowing Spectrum of Private, Protected Speech*, 99 CORNELL L. REV. 485, 485 (2014) (“Under the state-action doctrine, a plaintiff claiming a free-speech infringement must show some state action in order to trigger constitutional protection; the constraints of the First Amendment apply not to private persons but to the government.”).

240. *Id.*

241. *See supra* notes 112–45 and accompanying text

chill speech because these companies can decide that there is no value to having venomous language promoted on their platforms. It is also vital that companies are not complicit in this conduct, as users would be crippled without a platform where they can spread their ideology.

And yet, there are others who question the practicality of leaving “such . . . complex and momentous social decision[s] to the boards of . . . private corporations clustered in the San Francisco and Seattle metropolitan areas,” especially “[g]iven the relative absence of alternative channels of communication.”²⁴² Under this system, “the shape of free speech will be determined by popular opinion, market pressures, governmental pressures, and managerial conscience.”²⁴³

But given the Internet as a tool to educate and influence, and the expanding role of social media as the forum where this happens, compromises must be made to allow these websites to remain one step ahead of hate. Corporate suits deciding what speech they will allow on their services are not distinct from the private companies across the country who decide to terminate employees for the rhetoric they use in the office.²⁴⁴

Social networks should also log complaints to create a watch list. The networks could easily provide the watchlist to authorities should the need arise. They should also discourage people from posting hateful content by implementing a three-strike system, which would result in a ban and being added to the watch list.

B. Governmental Action

Essentially, any online hate speech regulations would have to pass the “strict scrutiny” test, meaning that they would have to be “narrowly tailored” to further a “compelling” governmental interest.²⁴⁵ A compelling interest is one that protects the public health, safety, or welfare.²⁴⁶ Drafted language

242. See Kyle Langvardt, *Regulating Online Content Moderation*, 106 GEO. L.J. 1353, 1385 (2018) (suggesting that content moderation could be left to the private sector but questioning whether corporations are an appropriate authority to regulate speech).

243. *Id.* at 1387.

244. According to the National Labor Relations Board, “the Board or an arbitrator must carefully balance” four factors when considering whether an employee’s workplace speech is protected: “(1) the place of the discussion; (2) the subject matter of the discussion; (3) the nature of the employee’s outburst; and (4) whether the outburst was, in any way, provoked by an employer’s unfair labor practice.” *Atlantic Steel Co.*, 245 N.L.R.B. 814, 816 (1979).

245. Winkler, *supra* note 69, at 800.

246. See *id.* (alteration in original) (footnote omitted) (“[T]he Court uses compelling in the vernacular to describe [the] societal importance’ of the government’s reasons for enacting the challenged law.”).

also must not be overbroad,²⁴⁷ meaning that it sweeps in too much speech, or vague,²⁴⁸ meaning that it is unclear as to what speech it sweeps under its umbrella. It is clear from the dramatic increase in hate crimes that this is an exceedingly important issue and one that the government should be obliged to address.²⁴⁹ Additionally, these non-speech elements that are secondary effects of hate speech should convince courts to find that speculative regulations pass constitutional muster.

Legislation or regulation could enhance the legal toolkit available to social media networks.²⁵⁰ As a baseline matter, Congress has the ability to regulate the Internet through the Commerce Clause, which gives it the authority to regulate “channels of interstate commerce,” “instrumentalities of interstate commerce,” and “activities having a substantial relation to interstate commerce.”²⁵¹ Various circuit courts have held that the Internet is a channel of interstate commerce.²⁵² If lawmakers were to introduce legislation on digital hate speech, they would have to set forth a definition of hate speech. South Africa has recently proposed a law criminalizing hate speech, which it defines as “any communication that intends to harm, incite harm or propagate hate on the basis of race, gender, religion or nationality, among other inherent characteristics.”²⁵³ First-time offenders could face three years in prison, while repeat offenders could be imprisoned for up to five years.²⁵⁴ Germany has passed a law that will fine social networks with over two million members up to fifty million euros if they do not remove “obviously illegal” posts up to twenty-four hours after they have been

247. See, e.g., *Gooding v. Wilson*, 405 U.S. 518, 530–31 (1972) (Burger, C.J., dissenting) (summarizing the purpose of the overbreadth doctrine).

248. See, e.g., *Nat’l Endowment for the Arts v. Finley*, 524 U.S. 569, 588 (1998) (discussing concerns of vagueness in statutory language).

249. See John Eligon, *Hate Crimes Increase for the Third Consecutive Year*, N.Y. TIMES (Nov. 13, 2018), <https://www.nytimes.com/2018/11/13/us/hate-crimes-fbi-2017.html> (reporting that “hate crime reports increased 17 percent” from 2016 to 2017 and that such “crimes remain vastly underreported”).

250. See *infra* notes 251–60 and accompanying text.

251. *United States v. Lopez*, 514 U.S. 549, 558–59 (1995) (citations omitted); U.S. CONST. art. I, § 8, cl. 3.

252. See, e.g., *United States v. Konn*, 634 F. App’x 818, 821 (2d Cir. 2015), *cert. denied*, 137 S. Ct. 149 (2016) (“[T]here can be no question that the Internet is a channel and instrumentality of interstate commerce” (citation omitted)); see also *United States v. MacEwan*, 445 F.3d 237, 245 (3d Cir. 2006) (“[T]he Internet is an instrumentality and channel of interstate commerce” (citation omitted)).

253. Krista Mahr, *A South African Law Would Make Hate Speech Punishable by Jail Time*, L.A. TIMES (Mar. 28, 2019), <https://www.latimes.com/world/la-fg-south-africa-hate-speech-20190328-story.html>.

254. *Id.*

informed about “law-breaking material” on their platforms.²⁵⁵ In order to ensure that any newly-drafted law passes constitutional muster, the language would have to be drafted in such a precise way that it could not be interpreted as overbroad or vague.²⁵⁶ There should also be a requirement that the speech targets a protected class.²⁵⁷

Regulations based in existing hate crime laws can also be drafted to expand protection into the digital realm. The Federal Communications Commission (FCC),²⁵⁸ a federal agency, could find that a clarifying regulation is necessary. This would also serve to define online hate speech and how it would be penalized. This process would allow for public involvement, as anyone could attempt to alter the proposed regulation by submitting a comment that the FCC would have to consider and respond to.²⁵⁹ In addition, regulations could create a uniform set of principles for social media websites regarding the need to monitor enumerated content and could impose monetary fines to websites that do not fully comply. Like sentence enhancements,²⁶⁰ regulations should also allow digital hate footprints to be used as evidence of any related crimes.

C. Global Action

European Union (EU) member governments have adopted a directive that strengthens regulations on video-sharing platforms and other newer forms of media, emphasizing the public interest.²⁶¹ The directive permits “governments to intervene in media activities on behalf of the public interest, as long as they are justified and proportionate.”²⁶² The new “measures take

255. *Germany Starts Enforcing Hate Speech Law*, BBC NEWS (Jan. 1, 2018), <https://www.bbc.com/news/technology-42510868>.

256. Mahr, *supra* note 253.

257. *Id.*

258. *What We Do*, FED. COMM’NS COMM’N, <https://www.fcc.gov/about-fcc/what-we-do> (“The Federal Communications Commission regulates interstate and international communications by radio, television, wire, satellite and cable in all 50 states, the District of Columbia and U.S. territories.”)

259. *See* Administrative Procedures Act, 5 U.S.C. § 553(c) (2018) (providing interested parties an opportunity to “participate in the rule making through submission of written data, views, or arguments” on proposed regulations).

260. *Sentence*, BLACK’S LAW DICTIONARY (11th ed. 2019) (an “enhanced sentence” is “[a] punishment more severe than usual, imposed typically because of the presence of aggravating circumstances”).

261. William New, *New EU Directive Limits Hate Speech, Establishes European Content Quotas*, INTELL. PROP. WATCH (June 11, 2018), <https://www.ip-watch.org/2018/11/06/new-eu-directive-limits-hate-speech-establishes-european-content-quotas/>.

262. *Id.*

aim at ‘incitement to violence or hatred,’ and ‘public provocation to commit a terrorist offence.’”²⁶³

Furthermore, “[t]o prevent and counter the spread of illegal hate speech online, . . . the [European] Commission agreed with Facebook, Microsoft, Twitter, and YouTube [to issue] a ‘Code of conduct on countering illegal hate speech online’ to help users notify[]” networks of “illegal hate speech” on their social platforms, and subsequently “improve the support to civil society as well as the coordination with national authorities.”²⁶⁴ Instagram, Google+, Snapchat, and other platforms later announced their intention to join the Code of Conduct as well.²⁶⁵

The right to freedom of expression is protected by Article 19 of the Universal Declaration of Human Rights (UDHR), an international treaty adopted by the United Nations General Assembly, and given legal force through Article 19 of the International Covenant on Civil and Political Rights (ICCPR).²⁶⁶ And “Article 20(2) of the ICCPR provides that any advocacy of national, racial, or religious hatred that constitutes incitement to discrimination, hostility, or violence must be prohibited by law.”²⁶⁷ A “[s]ix-factor threshold test” is recommended “to assist in judicial assessments of whether a speaker intends and is capable of having the effect of inciting their audience to violent or discriminatory action through the advocacy of discriminatory hatred.”²⁶⁸ The factors include: (1) the “context” (“the expression should be considered within the political, economic, and social context prevalent at the time it was communicated, for example the existence or history of conflict, existence or history of institutionalised discrimination, the legal framework, and the media landscape”); (2) the “[i]dentity of the speaker” (“the position of the speaker as it relates to their authority or influence over their audience, in particular if they are a politician, public official, religious or community leader”); (3) the “[i]ntent of the speaker to engage in advocacy to hatred; intent to target a protected group on the basis of a protected characteristic, and knowledge that their conduct will likely incite the audience to discrimination, hostility, or violence”; (4) the “[c]ontent of the expression” (“what was said, including the form and the

263. *Id.*

264. *Countering Illegal Hate Speech Online #NoPlace4Hate*, EUROPEAN COMM’N (Mar. 18, 2019), https://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=54300 (emphasis omitted).

265. *Id.*

266. ARTICLE 19, RESPONDING TO ‘HATE SPEECH’: COMPARATIVE OVERVIEW OF SIX EU COUNTRIES 8 (2018), https://www.article19.org/wp-content/uploads/2018/03/ECA-hate-speech-compilation-report_March-2018.pdf.

267. *Id.* at 9.

268. *Id.* at 10.

style of the expression, and what the audience understood by this”); (5) the “[e]xtent and magnitude of the expression” (“the public nature of the expression, the means of the expression, and the intensity or magnitude of the expression in terms of its frequency or volume”); and (6) the “[l]ikelihood of harm occurring, including its imminence” (“there must be a reasonable probability of discrimination, hostility, or violence occurring as a direct consequence of the incitement”).²⁶⁹

At the EU level, states are required to sanction racism and xenophobia through “effective, proportionate and dissuasive criminal penalties.”²⁷⁰ There are “four categories of incitement to violence or hatred offences that States are required to criminalise with penalties of up to three years,” but “States are afforded the discretion of choosing to punish only conduct which is carried out in ‘a manner likely to disturb public order’ or ‘which is threatening, abusive, or insulting.’”²⁷¹ “Criminal provisions directly restricting the most serious forms of ‘hate speech’ are provided in the criminal laws” of Austria, Germany, Hungary, Italy, Poland, and the United Kingdom, most evidently in the criminal or penal codes of those countries.²⁷²

The EU model can be influential in crafting regulations in the United States that would also permit our government “to intervene in media activities on behalf of the public interest.”²⁷³

D. Judicial Action

With the ease of using the Internet to disperse virulent messages, it is time for the Supreme Court to revisit its previous hate speech decisions to craft new law amenable to the modern age. Recently, a court decided that a neo-Nazi’s right to free speech did not extend to online harassment.²⁷⁴ On his website, the defendant “published over 30 related posts,” including “the phone numbers, email addresses and social media profiles of [the victim], her husband and 12-year-old son, as well as friends and colleagues.”²⁷⁵ Within a few months, the victim and her “family had received more than 700 vulgar and hateful messages, including death threats, many referencing the Holocaust. Some phone messages consisted solely of the sound of

269. *Id.* at 10–11.

270. *Id.* at 11 (citation omitted in original).

271. *Id.* at 11–12.

272. *Id.* at 15–16.

273. New, *supra* note 261.

274. Karen Zraick, *Neo-Nazis Have No First Amendment Right to Harassment, Judge Rules*, N.Y. TIMES (Nov. 15, 2018), <https://www.nytimes.com/2018/11/15/us/daily-stormer-anti-semitic-lawsuit.html>.

275. *Id.*

gunshots.”²⁷⁶ The judge denied a motion to dismiss on First Amendment grounds and wrote that the victim was a private citizen, not a public figure, and that the defendant incited his followers to harass her as part of a personal campaign.²⁷⁷

A white supremacist was sentenced to 42 months in prison for soliciting violence against a jury foreman who convicted another white supremacist for “soliciting the murder of a federal judge.”²⁷⁸ On his website, he instructed others to injure the juror and posted derogatory comments about him.²⁷⁹ Prior to this incident, the same defendant posted to his site “the home address[es] and/or personal identifying information of individuals who were targets of criticism on the Internet.”²⁸⁰ He was also aware that “the target audience of his web site . . . at times engaged in acts of violence, directed at non-whites, Jews, gays and persons perceived by white-supremacists as acting contrary to their interests.”²⁸¹

Thus, it is incumbent upon the Supreme Court to address the convergence of hate speech and physical violence or the threat of physical violence, and to construct a new framework for when this speech is unconstitutional and thus proscribable.

E. Oversight Review Board

Perhaps the most advantageous solution is to create a review board, consisting of members from various spaces, including the government (both on federal and state levels), federal agencies, academia, journalism, the law, psychology, and educators. Such a board would be responsible for oversight functions, establishing uniformity, and compliance. It would also put the onus on a group of experts from a diversity of industries, thereby assuring that no single population has absolute power in policing and punishing hate speech. The board could draft a definition of hate speech as well as policies for social media platforms. This would place all sites on equal footing and provide users with a bright-line set of expectations as to what they can and cannot do online.

276. *Id.*

277. *Id.*

278. *White-Supremacist William White Sentenced to 42 Months in Prison for Soliciting Violence Against Hale Jury Foreman*, U.S. DEP’T OF JUSTICE (Feb. 20, 2013), <https://www.justice.gov/usao-ndil/pr/white-supremacist-william-white-sentenced-42-months-prison-soliciting-violence-against>.

279. *Id.*

280. *Id.*

281. *Id.*

If any governmental actors were involved with the board, there would be state action, meaning that the First Amendment would be implicated.²⁸² Therefore, strict scrutiny would have to be met, and there could be neither subject matter nor viewpoint discrimination.²⁸³ Consequently, the board could focus heavily on the secondary effects of hate speech,²⁸⁴ as the Court has explained that when the government can recognize secondary effects, the regulation is “*justified* without reference to the content of the . . . speech.”²⁸⁵ If the government can point to an increase in crime as a secondary effect that it seeks to prevent by regulating hate speech, then a court may treat the regulation as content-neutral even though it is content-based.

Finally, there will be those who will claim that opponents of hate speech should focus on countering that speech.²⁸⁶ But the theory of counter-speech is inapposite in the debate over hate speech.²⁸⁷ Hate cannot drive out hate. The lynchpin is that hate speech does not contribute to the “marketplace of ideas.”²⁸⁸ Justice Breyer has suggested that:

whenever government disfavors one kind of speech, it places that speech at a disadvantage, potentially interfering with the free marketplace of ideas and with an individual’s ability to express thoughts and ideas that can help that individual determine the kind of society in which he wishes to live, help shape that society, and help define his place within it.²⁸⁹

While it is true that a tenet of our Constitution and our nation is the ability to speak freely and autonomously to formulate our own views and

282. Wirth, *supra* note 239, at 485.

283. *Turner Broadcasting System, Inc. v. F.C.C.*, 520 U.S. 180, 234 (1997).

284. *See, e.g., Renton v. Playtime Theatres*, 475 U.S. 41, 47 (1986) (emphasis omitted) (distinguishing between concerns about the content of adult movie theatres and the “secondary effects of such theaters on the surrounding community”).

285. *R.A.V. v. St. Paul*, 505 U.S. 377, 389 (2012) (emphasis in original) (omission in original) (quoting *Renton*, 475 U.S. at 48).

286. Nadine Strossen, *Don’t Silence Graduation Speakers: Fight Hate Speech with More Speech*, USA TODAY (May 1, 2018), <https://www.usatoday.com/story/opinion/2018/05/01/censorship-hate-speech-freedom-first-amendment-column/564868002/>.

287. *Whitney v. California*, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring) (“[T]he remedy to be applied is more speech, not enforced silence.”).

288. Richard Delgado & Jean Stefanie, *Images of the Outsider in American Law and Culture: Can Free Expression Remedy Systemic Social Ills*, CORNELL L. REV. 1258, 1260 (1992) (discussing the fallacy that the “marketplace of ideas” remedies phenomena such as hate speech and hateful depiction because of temporal blindness).

289. *Reed v. Town of Gilbert*, 135 S. Ct. 2218, 2234 (2015) (Breyer, J., concurring).

opinions, that freedom is not limitless.²⁹⁰ It is important to remember that the Court has explained that, “[i]t has been well observed that [fighting words] are no essential part of any exposition of ideas, and are of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality.”²⁹¹ One purpose of free speech is to permit people to judge the truth and acceptability of various sentiments.²⁹² But hate speech is not grounded in veracity, and it should not be universally perceived as acceptable. Speech must have value to warrant protection, and, as evinced by multiple instances occurring on local to global scales, hateful speech is dramatically leading to acts of violence, whether by the perpetrators of such words or by their voracious followers.²⁹³ When malevolent words are leading to murderous actions, a paramount question that society must ask of itself is whether that value exists. Only then can such rhetoric be properly regulated.

CONCLUSION

As the wave of a very dangerous social climate crashes upon our nation’s shores, it is high time to regulate hate speech online. Hate speech pumps new blood through the veins of hate groups and invigorates them to take the rhetoric that is dispersed online out to the streets to target minority groups and those of protected classes. Now, people no longer have to stand on soapboxes in the village square or hand out leaflets in order to broadcast a caustic message of hate or intolerance. The accessibility by which we have the opportunity to send and receive vitriolic messages is quite literally at our fingertips and can make it seem as though this speech is ubiquitous. Coupled with the boundless nature of cyberspace, these rancorous words are proliferated at lightning fast speed. The Holocaust, one of the largest mass genocides of all time, was not started by violence. It was started by words.²⁹⁴ It was started by a group of people who shared a common belief that some people were more worthy than others because of uncontrollable qualities. This is not a political issue, a cultural issue, or a partisan issue—it is a basic human rights issue. It is about right versus wrong, with one clear-cut answer:

290. See *Chaplinsky v. New Hampshire*, 315 U.S. 568, 571 (1942) (footnote omitted) (“[I]t is well understood that the right of free speech is not absolute at all times and under all circumstances.”).

291. *Id.* at 572 (footnote omitted).

292. Schultz, *supra* note 17.

293. *Hate Speech on Social Media: Global Comparisons*, COUNCIL ON FOREIGN RELATIONS (June 7, 2018), <https://www.cfr.org/backgrounder/hate-speech-social-media-global-comparisons>.

294. *Museum Condemns White Nationalist Conference Rhetoric*, U.S. HOLOCAUST MEM’L MUSEUM (Nov. 21, 2016), <https://www.ushmm.org/information/press/press-releases/museum-condemns-white-nationalist-conference-rhetoric>.

not to divide or single out people because of their inherent qualities. Instead, as a society, it is incumbent that we come together to drive out such hate and promote understanding, tolerance, and unification.

