2017

Hate Speech on the Internet: Crime or Free Speech?

Grace Udoh-Oshin
Long Island University, grace.udoh-oshin@my.liu.edu

Hate Speech on the Internet: Crime or Free Speech?

An Honors College Thesis

By

Grace Udoh-Oshin

Spring, 2017

Department of Communications and Film

_____

Advisor: Professor Carolyn Schurr Levin

_____

Reader: Professor Benjamin Gerdes

4/28/17

_____

Date

Abstract

The Internet has a lot of positive features. It gives users fast access to information, allows for communication with others who might otherwise be out of reach, and much more; however, the Internet allows for a lot of negativity as well. There has been an increase in hate based activity on social media, and the anonymity and flexibility afforded by the Internet has made harassment and expressions of hate easy, thus making it much harder to implement traditional law enforcement. This paper gives a history of hate speech in America, and examines the difficulties of regulating hate speech on the Internet through various hate speech cases. It explores the need of the legal system to clarify the difference between unprotected and protected speech, and narrow and clarify what constitutes a "true threat" or "intent". The paper progresses to consider how social media platforms stop harm caused by hate speech and includes solutions proposed by legislators on how to regulate unprotected categories of speech. Further, it argues that there needs to be a clear, universally accepted definition of hate speech, and provide lower Courts with guidance. This paper proves that without a proper clarification or a definition, it will remain confusing for all.

TABLE OF CONTENTS

ABSTRACT 2

INTRODUCTION 4

THESIS/PREVIEW 5-6

MAIN POINT 1 6

HISTORY OF HATE SPEECH 7
HATE SPEECH DEFINED 8
RELEVANCY OF TESTS 9-13

MAIN POINT 2 13

TRUE THREATS/INCITEMENT TO IMMINENT LAWLESS ACTION 15
CYBERBULLYING 26

MAIN POINT 3 34

DOES THE INTERNET ENABLE HATE SPEECH 34
ARGUMENTS FOR AND AGAINST REGULATING HATE SPEECH 41

MAIN POINT 4 47

MAIN POINT 5 55

CONCLUSION 57

BIBLIOGRAPHY 58-61

Introduction:

On December 8, 2010, Anthony Elonis was arrested and charged with five counts of violating a federal anti-threat statute. Elonis had posted statements on his Facebook page that appeared to threaten his ex-wife, a kindergarten class, the local police, and an FBI agent. Despite the fact that his ex-wife and others perceived his posts as threats, Elonis argued that he didn't mean what he said in a literal sense. He claimed he was an aspiring rap artist and that his comments were simply a form of artistic expression. According to Adam Liptak, who covers the United States Supreme Court and writes a column on legal developments called "Sidebar." Elonis wrote in his FaceBook posts:

> "He would like to see a Halloween costume that included his wife's "head on a stick." He talked about "making a name for myself" with a school shooting, saying, "Hell hath no fury like a crazy man in a kindergarten class." He fantasized about killing an F.B.I. agent. "Pull my knife, flick my wrist, and slit her throat," he wrote." (Liptak, 2015)

Anthony D. Elonis was convicted under 18 U.S.C. § 875(c), a federal statute that prohibits making "any threat to injure the person of another" via the Internet. While some would argue that the posts were threatening in nature, the Supreme Court sided with Elonis, and on June 1, 2015, in an 8-1 decision, it reversed his conviction stating that prosecutors did not do enough to prove Anthony Elonis's intent when he published the threats on his Facebook account (*Elonis v. United States*, 2015). Although Elonis' rants on social media were disturbing, the Supreme Court held that the speech was protected by the First Amendment because the speech did not

intend to threaten or deliberately harm. However, the First Amendment protection of speech does not distinguish between lawful and unlawful expressions. Essentially, the Court maintained that hate speech in social media is protected by the Constitution, even though there is ambiguity to what exactly constitutes a true threat.

According to Jeremy Waldron, a legal philosopher, and author of the book, "The Harm in Hate Speech", "I think it is unlikely that legislation of the kind [laws prohibiting hate speech] will ever pass constitutional muster in America" (Waldron 11). But maybe it is time to at least define and narrow the type of speech that is unprotected in America. This is particularly significant as hate speech is increasingly targeting celebrities, and public figures, many of whom depend on social media to connect with their fans and promote themselves and their work. Some people have argued that the growing vitriol and hate speech directed at celebrities, and different minority groups--blacks, immigrants, the LGBT community and women--is dangerous, threatening, and undermines public good. However, others argue that any effort to curtail hate speech on social media is a clear violation of their First Amendment right to free speech.

Thesis Statement/purpose:

The Anthony Elonis case and others like it highlight a problem, which is that there are inconsistencies in the law regarding what is considered protected speech on the Internet. Recent developments and intensity of hate speech online have heightened the need for some sort of regulation to differentiate protected speech and speech that is not protected by the First Amendment on the Internet. While it could be argued that perhaps the law has not yet caught up with the rapidly changing technology of the Internet, nevertheless, this inconsistency raises an important question: When does free speech cross over to non-protected speech that is considered

a "true threat" on the Internet, especially on social media; specifically, Twitter, Facebook, Reddit, and Youtube? This thesis argues that the Supreme Court has ruled that hate speech without the intent of violence is protected on the Internet, but perhaps it is time for the law to (1) clarify the difference between unprotected and protected speech, (2) narrow and clarify what constitutes a "true threat" (3) create a universally accepted definition of hate speech; and (4) draw the line between protected speech and cyberbullying on the Internet, because currently that line is blurred.

Preview:

This research paper seeks to examine and analyze hate speech on the Internet, specifically on Twitter and Facebook by answering the following questions:

1) What is the history of hate speech under the First Amendment? How is hate speech defined? What are the tests used to decide a hate speech case, and are they still relevant in today's society?
2) When is hate speech not protected by the First Amendment? ("true threat" cyberbullying statutes, incitement)
3) Does the Internet, including social media allow for anonymity, and a broader audience thus enabling hate speech? What are the arguments for and against regulating hate speech on social media?
4) What implications are there, if any, for social media platforms such as Twitter and Facebook when it comes to hate speech?
5) Is there a solution that will minimize hate speech on social media, while finding ways to not invade individuals' First Amendment rights?

<u>Main Point # 1:</u> What is the history of Hate Speech under the First Amendment? What is hate speech, and how does the law define it? What are the tests used to decide a hate speech case, and are those tests still relevant in today's society?

History

Freedom of Speech is a cherished liberty that protects and allows even the most revolting statements. According to the U.S. Supreme Courts' interpretation of the First Amendment; protected speech includes hate speech, symbolic speech, protests, and written or typed speech. As an American you have the right to say just about whatever you want without fear of punishment by the government. Since the Bill of Rights was added to the U.S. Constitution in 1791, freedom of speech has been a staple of American liberty. According to M. Alison Kibler, an associate professor of American Studies and Women's and Gender Studies at Franklin and Marshall College in Pennsylvania, the history of hate speech stems from racial tensions affecting American minorities:

> "At the turn of the twentieth century, states and municipalities throughout the country banned "racial ridicule" on stage and screen. Widely deemed to be separate white races in the early twentieth century, Irish and Jewish citizens joined African Americans in their support for the censorship of racist images in advertising, theater and motion pictures. Jewish and African Americans were alarmed that images of their respective groups as sexually depraved or criminal contributed to vigilante violence against them, although the extent of that threat was much greater for African Americans. All three groups believed that disparaging representations impaired their social standing and political equality" (Kibler, 2015).

However, dealing with hate speech has been problematic throughout American history because two deeply rooted American values, free speech and equality, are constantly challenged

(Kibler, 2015). Legislations have shifted and terms and restrictions have changed over time. Kibler wrote:

> "The long history of hate speech legislation shows that popular entertainment has been a rich arena for the regulation of racial ridicule in the early twentieth century…lawmakers and activists on the left and the right evaluated the harms of particular speech within varied contexts to craft compromises between racial equality and freedom of expression." (Kibler, 2015)

Hate Speech Defined

Hate speech is speech that offends, threatens, or insults groups, based on race, color, religion, national origin, sexual orientation, disability, or other traits (*Debating the Mighty Constitutional Opposites*). Examples of hate speech can include racist cartoons, anti-Semitic symbols, ethnic slurs or other derogatory labels for a group, burning a cross, politically incorrect jokes, sexist statements, anti-gay protests...etc. Although we see a lot of hate speech online, hate speech is not new and neither is the goal of hate speech. According to an article written by Sean McElwee, a researcher and writer based in New York:

> "Hate speech aims at two goals. First, it is an attempt to tell bigots that they are not alone. Frank Collins — the neo-Nazi prosecuted in National Socialist Party of America v Skokie (1977) — said, "We want to reach the good people, get the fierce anti-Semites who have to live among the Jews to come out of the woodwork and stand up for themselves." The second purpose of hate speech is to intimidate the targeted minority, leading them to question whether their dignity and social status is secure." (McElwee,

2013).

Relevancy of Tests

Through the outcome of several court cases, the Supreme Court has maintained that hate speech is fully protected by the First Amendment; regardless of how it is communicated as long as the speech does not incite violence. However, there was a time in which the government had wide-ranging influence in prosecuting those who engaged in speech advocating violence or legal activity; but only if the speech incited violence or created a clear and present danger. For example, in the case *Schneck V. United States* (1919), Charles Shneck mailed pamphlets suggesting that the draft was wrong and the war was motivated by capitalist greed, during World War I. The pamphlets read "Do not submit to intimidation", and advised peaceful actions against the government, such as petitioning. Schenck was charged with conspiracy to violate the Espionage Act by attempting to cause mutiny in the military and to hinder recruitment. (Schenck V. United States, 249 U.S. 47, 1919). In the case of *Schneck V. United States*, the court contended that:

> "Schenck is not protected in this situation. The character of every act depends on the circumstances. "The 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." During wartime, utterances tolerable in peacetime can be punished." (*Schenck v. United States*, 249 U.S. 47).

This landmark case created an opportunity for the government to get past the U.S. Constitution's

strong protections for speech. According to Alex McBride, a third year law student at Tulane Law School in New Orleans and articles editor of the Tulane Law Review:

> "The Supreme Court invented the famous "clear and present danger" test to determine when a state could constitutionally limit an individual's free speech rights under the First Amendment. In reviewing the conviction of a man charged with distributing provocative flyers to draftees of World War I, the Court asserted that, in certain contexts, words can create a "clear and present danger" that Congress may constitutionally prohibit. While the ruling has since been overturned, Schenck is still significant for creating the context-based balancing tests used in reviewing freedom of speech challenges." (McBride, 2006)

Since *Schenck v. United States*, the clear and present danger test has been applied in a multitude of older cases; Abrams v. United States (1919), Whitney v. California (1927) and Dennis v. United States (1951). Again, the Government was able to surpass Freedom of Speech protection when in 1948, the leaders of the Communist Party of America were charged with violating provisions of the Smith Act. The Act made it unlawful to knowingly conspire to teach and advocate the overthrow or destruction of the United States government. In a 6-to-2 decision, the Court maintained that the Communist Party were in violation of the First Amendment, and that the Smith Act did not violate it. The court contended that their actions created a "clear and present danger" to the U.S. government. (*Dennis v. United States*, 341 U.S. 494). This case introduced a "balancing test" allowing the Supreme Court to assess free speech challenges against the state's interests on a case-by-case basis.

Things changed in 1969 with the *Brandenburg v. Ohio* case. Another landmark case that is significant, and well-known for replacing the Clear and Present danger test with the two part-test that determined whether speech would provoke imminent lawless action. In 1969, a prominent leader of the KKK made a violent speech about getting revenge against Blacks and Jews, which violated an Ohio criminal law that prohibits public speech that advocates illegal activities. It is important to note that during the Klan rally in Ohio, there were also men carrying firearms. Although the words were violent, it was maintained by the Supreme Court that his words did not provoke imminent lawless action, and that the Ohio justice system violated Brandenburg's right to free speech. The Court created three distinct elements of this test: intent, imminence, and likelihood. (1) the speech is "directed to inciting or producing imminent lawless action," and (2) the speech is also "likely to incite or produce such action. (*Brandenburg v. Ohio*, 1969). The Supreme 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." (*Brandenburg v. Ohio*, 1969).

This case highlights conflicts the court has with distinguishing protected and unprotected speech. The Clear and Present danger test had to be modified and adapted by a newer test in order to find a balance between what is considered protected and when speech becomes "threatening". In the case of *Schenck v. United States*, Schneck was convicted even though his actions did not cause violence or illegal activity, not was there any imminent threat. The same point was made in

Dennis v. United States:

> "The criminal syndicalism act made illegal the advocacy and teaching of doctrines while ignoring whether or not that advocacy and teaching would actually incite imminent lawless action. The failure to make this distinction rendered the law overly broad and in violation of the Constitution." (*Dennis v. United States*, 1951)

Both defendants in those cases were prosecuted even though today their speech would have been protected following current standards. Although both cases were later reversed, it took until *Brandenburg v. Ohio* for a change to be made, and for the Supreme Court to realize how broad these tests were, and still are. Almost half a century after Brandenburg V. Ohio, this test is still used today, along with the more frequently used balancing test; and neither tests have yet to define what constitutes a true threat. Since *Brandburg v. Ohio*, many hate speech cases have been ruled in the defendant's favor. While online speech content delivery differs from symbolic or spoken speech delivery, the Supreme Court uses the same tests across all platforms of speech in deciding whether the speech is protected or unprotected. This author argues that it is time for these tests to be updated to better suit the technology driven world in which we live. According to First Amendment expert and law professor David L. Hudson Jr.,

> "Some legal commentators have urged that the Brandenburg standard be modified with respect to online hate speech. One commentator wrote…"New standards are needed to address the growing plague of Internet speech that plants the seeds of hatred, by combining information and incitement that ultimately enables others to commit

> violence." Another agreed, writing: "Although Brandenburg may be suitable for the traditional media outlets, which were well-established when it was decided, Internet speech and many unforeseen changes have made such a standard outdated." (Hudson, 2002)

Instead of the Courts continuing to adapt the Brandenburg test, I maintain that it is time to stop using outdated standards and just create entirely new ones. Hate speech is fully protected under the First Amendment as long as it does not **incite** violence, but what about speech that does, and is therefore **not** protected by the First Amendment.

<u>Main Point # 2:</u> When is hate speech not protected by the First Amendment (cyberbullying statutes, "true threat", incitement to imminent lawless action)?

Although most speech is protected under the First Amendment, there are a few classification of speech that are not protected such as:

- Obscenity -- i.e., erotic expression, grossly or patently offensive to an average person, that lacks serious artistic or social value -- may be prohibited.
- Child pornography may be banned whether or not it is legally obscene and whether or not it has serious artistic or social value, because it induces people to engage in lewd displays, and the creation of it threatens the welfare of children.
- Defamatory statements may be prohibited. (In other words, the making of such statements may constitutionally give rise to civil liability.) However, if the target of the defamation is a "public figure," she/he must prove that the defendant acted with

"malice." If the target is not a "public figure" but the statement involved a matter of "public concern," the plaintiff must prove that the defendant acted with negligence concerning its falsity.

- Commercial Speech may be banned only if it is misleading, pertains to illegal products, or directly advances a substantial state interest with a degree of suppression no greater than is reasonably necessary." (Fisher, 2001)

Most categories of hate speech are protected under the First Amendment unless, according to Justice Frank Murphy, who ruled in the *Chaplinsky v. New Hampshire* case.

"There are certain well-defined and limited classes of speech, the prevention and punishment of which have never been thought to raise a Constitutional problem. These include the lewd and obscene, the profane, the libelous and the insulting or 'fighting' words – those which by their very utterances inflict injury or tend to incite an immediate breach of the peace." (Chaplinsky v. New Hampshire, 1942)

Today, hate speech cases are decided based on old tests set forth by the Supreme Court which would prove if there is enough substantial justification for the speech to be ruled unprotected. Furthermore,

"Content-based governmental restrictions on speech are unconstitutional unless they advance a "compelling state interest." This allows speech to be banned only when it is intended—and likely—to incite imminent violence or lawless action." (Fisher, 2001)

True Threats/Incitement to imminent lawless action

If speech is considered to be a true threat or the speech will lead to imminent lawless action which are categorized under "fighting words," defined as words which by their very utterance inflict injury or tend to incite an immediate breach of the peace; the speech is not protected under the First Amendment. Nevertheless, what exactly constitutes a true threat? This is where the ambiguity lies, and according to Hudson Jr.,

> "The First Amendment does not give a person the right to walk up to someone else and say "I am going to kill you" or to announce in an airport, "I am going to bomb this plane." Yet the line between protected expression and an unprotected true threat is often hazy and uncertain. What if a speaker makes a seemingly threatening statement about a political figure through the use of hyperbole? What if a student says that if he receives a poor grade, he may "go Columbine"? What if an abortion protester talks about participating in a "war against abortionists"? (Hudson, 2008)

Hudson's point brings attention to the fact that whatever constitutes a "true threat" needs to be clarified. The U.S. Supreme Court tackled a prominent true-threat case in *Watts v. United States* (1969). In this case an African-American man, Robert Watts, was arrested and prosecuted for making threating allegations to the U.S. president while protesting in Washington D.C.:

> "They always holler at us to get an education. And now I have already received my draft classification as 1-A and I have got to report for my physical this Monday morning. I am not going. If they ever make me carry a rifle, the first man I want to get in my sights is

> L.B.J. They are not going to make me kill my black brothers." (*Watts v. United States*,1969)

Watts was charged with violating a federal law that prohibits threats against the President; however, he appealed to the Supreme Court, which reversed the ruling stating that Watts' statement was political hyperbole instead of a true threat.

> "We agree with [Watts] that his only offense here was 'a kind of very crude offensive method of stating a political opposition to the President,'" the Court wrote in a per curiam opinion [a decision delivered via an opinion issued in the name of the Court rather than specific judges]. "Taken in context, and regarding the expressly conditional nature of the statement and the reaction of the listeners, we do not see how it could be interpreted otherwise." (*Watts v. United States*, 1969)

However, the Supreme Court made its ruling in the case without defining what constitutes a true threat. The Court's decision was based on the context of the speech, not the speech itself. Lower Courts considering true-threat cases have focused on certain elements of *Watts v. United States,* including: (1) the fact that the comments were made accompanying a political debate; (2) the conditional nature of the threat; and (3) the context of the speech, as apparently several listeners laughed after Watts spoke. Since *Watts v. United States*, Courts have allowed the use of an objective standard for determining a speaker's intent rather than a subjective standard. The government no longer needs to prove that the speaker's words intended to threaten, but rather that a reasonable person would have perceived the words as a threat.

(Dussich, 2015). The First Amendment takes the preferred position in the balancing test. However it is difficult to prove what an individual intended when making a verbal threat, and most Courts sided with the speakers in hate speech cases such as *Chaplinsky v. New Hampshire*, *Watts v. United States*, and *R.A.V. v. City of St. Paul* . The First Amendment takes the preferred position in the balancing test used by the Supreme Court.

Though, years later in *Virginia v. Black*, the Supreme Court partially agreed with the state. Barry Black along with other individuals were convicted individually for burning crosses in public and private properties, which violated a Virginia statute that makes it a felony "for any person..., with the intent of intimidating any person or group..., to burn...a cross on the property of another, a highway or other public place." It also specified that "any such burning...shall be prima facie evidence of an intent to intimidate a person or group." Black was found guilty at his trial and the Virginia Supreme Court held that, "a state, consistent with the First Amendment, may ban cross burning carried out with the intent to intimidate." In so doing, the Court considered the speech to be constitutionally unprotected "true threats." Under that section, "a State may choose to prohibit only those forms of intimidation that are most likely to inspire fear of bodily harm." *(Virginia v. Black,* 2003)

The Virginia State Supreme Court reversed the convictions, ruling that the statute was a violation of First Amendment rights and unconstitutionally overbroad because of the intent standard. However, The United States Supreme Court affirmed the state court's decision to overturn the conviction of the Ku Klux Klan cross-burners. The Court found that the broader purpose of the prohibiting cross-burning statute was acceptable under the First Amendment, but

that it was otherwise unconstitutional. However, it is clear that the two ideas are contradicting. The Court disagreed with a part in Virginia's statute which stated, "Any such burning of a cross shall be *prima facie* evidence of an intent to intimidate a person or group of persons," ruling that it was unconstitutional because of its "indiscriminate coverage." The state must prove intent to intimidate. *(Virginia v. Black,* 2003)

Supreme Court Justice David Souter argued in dissent that cross-burning, even with the proven intent to intimidate, should not be a crime under the R.A.V. v. City of St. Paul precedent because of "the statute's content-based distinction." (*Virginia v. Black,* 2003). Justice Scouter's dissent suggests that there is confusion among the court when deciding what kind of speech is unprotected.

Flash forward to present day and the Courts have yet to define this category of unprotected speech. In the recent case of *Elonis v. United States*, the Supreme Court held that Elonis's social media postings were protected by the First Amendment due to his speech not being considered to be a true threat, or having intent because he didn't really mean what he said. This brings up an interesting question: how can a court determine the intent of the speech, and should they be able to? The issue being, "does a conviction of threatening another person under federal anti-threat statute 18 U.S.C. § 875(c) require proof that the defendant meant what he said in a literal sense?" (*Elonis v. United States*, 2015). According to the Harvard Law Review,

> "Two central and unresolved questions are first whether a defendant can be convicted under the federal threats statute absent proof that he subjectively intended to threaten

> anyone, and second, if the statute itself does not require this evidence, whether the First Amendment does." ("Harvard Law Review: *Elonis v. United States*").

Instead of coming up with a definition, Courts have cited that if a "reasonable person" understands the defendant's words to be a threat then the defendant will assume criminal liability. However according to the Harvard Law Review:

> "Last Term, in *Elonis v. United States*, the Supreme Court disagreed with this standard and held that a conviction under 18 U.S.C. § 875(c) may not be based solely on a reasonable person's interpretation of the defendant's words.*7. Id. at 2012–13.* The Court reversed the defendant's conviction on narrow statutory grounds rooted in principles of mens rea (guilty mind – criminal intent) and did not reference the constitutional "true threats" doctrine. The majority, however, left undecided the minimum mental state required for criminal liability. As a result, lower Courts are left to answer both questions originally presented in *Elonis*. Certain parallels between obscenity and threats, hinted at in *Elonis*, suggest one approach to this task." ("Harvard Law Review*: Elonis v. United States*")

Whether the speech is spoken or typed, the Courts have indicated that the First Amendment does not protect "true threats". Throughout legal history, many cases have highlighted the elements required to define the validity of a threat, however criteria change from case to case and without a clear definition Courts are left to decide based on what they have come to consider constitutes a "true threat." *(Determining the Legality of a Threat*). Even though

Elonis's speech was protected under the First Amendment, other forms of hate speech communicated online could fall into the unprotected category of a "true threat", depending on intent found. According to First Amendment expert, Hudson Jr.:

> "The First Amendment does not protect an individual who posts online "I am going to kill you" about a specific individual. A few cases have applied the true-threat standard to online speech. *In Planned Parenthood v. American Coalition of Life Activists* (2002), the 9th U.S. Circuit Court of Appeals held that some vigorous anti-abortion speech — including a Web site called the *Nuremberg Files* that listed the names and addresses of abortion providers who should be tried for "crimes against humanity" — could qualify as a true threat." (Hudson, 2002)

Hudson's statement highlights a problem that arises without a clarified understanding of what constitutes a true threat. Hudson states that the First Amendment does not protect an individual who posts online, "I am going to kill you," however Elonis posted on Facebook, "there's one way to love you but a thousand ways to kill you". Both are threating statements, but are written differently. According to Rebecca Dussich, Associate Member, *University of Cincinnati Law Review:*

> "Elonis does not dispute that he posted Facebook status messages regarding his desire to kill his wife, detonate bombs in the presence of law enforcement, and shoot up a local elementary school (among other threats). Rather, he disputes that these were intended as threats, stating that he was merely 'expressing frustration'." (Dussich, 2015)

One could argue that the posts have different contextual meanings, but regardless of how both posts were communicated, they were still considered threats to his ex-wife, and others affected by his posts. In reference to the *Virginia v. Black* case, there is a confusing difference between cross-burning seen as an intimidation tactic, which was found not to be in violation of the First Amendment citing intent to threaten was clearly there. When compared to the *Elonis v. United States* case, in which the First Amendment protected his threatening words because they were not deemed "true threats," it noticeable that the polarity in both cases verify the need for a unified law to clarify the difference between protected and unprotected speech. Although there is a distinction between symbolic speech, plus conduct and online hate speech, this is even more of a reason for Courts to consider the inconsistencies in regulations of speech. According to Sherry F. Colb, a Justia (a website that provides free case law, codes, regulations and legal information for lawyers, business, students and consumers world wide) columnist, and Professor of Law and Charles Evans Hughes Scholar at Cornell University;

> "A true threat is a threatening utterance, i.e., a statement that is not just hyperbole, not just a joke, and not a comparably non-threatening threat-like statement…True threats are subject to criminal prohibition because they have harmful effects on their targets, even when the threats are not actually carried out. If someone were to call John Doe on the telephone and say that he plans to blow up Doe's child's school today, then Doe would probably feel extremely scared, keep his child home from school, and maybe also notify the school office of the threat and thereby motivate the school to close for the day, pending a bomb squad clearance of the school grounds. Even if the caller has no plans of setting off any explosives, then, the result of his "speech" is to generate terror, to keep

children out of school for a day, and to cost both money and attentional resources to ensure the safety of the school. All of these effects are very destructive…" (Colb, 2014)

Elonis posted "rap lyrics" on Facebook threatening to shoot up a school, and although others do not know his intentions, Elonis's post raised some red flags. Elonis posted these "rap lyrics" on his Facebook page:

I'm checking out and making a name for myself.
Enough elementary schools in a ten-mile radius
To initiate the most heinous school shooting ever imagined.
And hell hath no fury like a crazy man in a Kindergarten class
The only question is … which one? (*Elonis v. United States,* 2015).

According to Colb, "One example of something that would not count as a true threat would be the Elvis Costello song "My Aim is True," which might sound, out of context, like a threat to kill a woman named Allison, but is in fact just a song." (Colb, 2014). Elonis argued that like Elvis Costello's song, his posts were just creative rap lyrics taken out of context. Now, had Elonis posted the status on Facebook, tagging an elementary school directly and wrote, "I am going to shoot the school", according to Colb's definition, it could be considered a true threat. His comments could cause a harmful effect on his targets.

However, if a person posted a status to Facebook directly saying, "I am going to kill you" that could also be argued as "expressing frustration". The problem that the Courts face is, "since

determining when lawful speech over steps the boundaries and becomes unlawful is so difficult, the definition of a "true threat" is open to interpretation. The definition is, therefore, continuously redefined, resulting in unbalanced judicial hearings" (*Determining the Legality of a Threat*). Had Elonis actually followed through with his threats then he would have been convicted for his conduct, but you cannot bring back all the peoples whose lives would have been lost. With that being said, it could be time for the Judicial System to start limiting more speech on the Internet. According to the First Amendment Center website:

> "Some Courts have determined that "if a reasonable person would foresee that an objective rational recipient of the statement would interpret its language to constitute a serious expression…[then] the message conveys a 'true threat'. Other Courts consider a series of factors in determining whether speech constitutes a true threat, including (1) the reaction of the recipient of the speech; (2) whether the threat was conditional; (3) whether the speaker communicated the speech directly to the recipient; (4) whether the speaker had made similar statements in the past; and (5) whether the recipient had reason to believe the speaker could engage in violence (*How do Courts determine whether speech is a true threat*).

In the Elonis case, Anthony Elonis "claimed that his words were "constitutionally protected works of art," however, context revealed them to be unprotected "real threat[s] in the guise of rap lyrics" ("Harvard Law Review: *Elonis v. United States*"). This presents an issue for Courts and jurors who may have differing views on the context of the defendant's speech. According to the Harvard Law Review:

> "Courts must now ascertain whether a reckless mental state adequately separates "wrongful" from "innocent" conduct in the context of threatening speech…In fact, the reasoning in many "true threats" cases still looks a lot like this: "I know it when I see it. But if criminal liability turns solely on whether the defendant subjectively knew of and disregarded the risk that his message was objectively a threat, then clearly delineating ex ante what constitutes a constitutionally unprotected threat is paramount." ("Harvard Law Review: *Elonis v. United States*").

If the lower Courts don't receive guidance from the U.S. Supreme Court about what type of speech is considered a true threat, then how will they equally, and fairly across the board decide what type of threating speech is protected and not protected under the First Amendment. With the advent of the digital age and hate speech growing online, the Courts have yet to come up with a definition. How can Courts determine the intent of speech when Courts have differing views, and if there is no universally updated test? According to a 2015 Times magazine article by Soraya Chemaly, the director of the Women's Media Center Speech Project and organizer of Safety and Free Speech Coalition; and Mary Anne Franks, a law professor at the University of Miami School of Law and the Vice-President of the Cyber Civil Rights Initiative:

> "The court's narrow decision provides little guidance to Courts struggling with the issues raised by threatening speech—online or offline—and raises troubling issues for victims of threats, especially in the context of domestic violence." (Chemaly and Franks, 2015)

If the U.S Supreme Court and lower Courts were able to create and follow complete, narrowed tests, then the outcome of hate speech cases would be justified with reasonable evidence and support, citing tests that clarify what a "true threat" is. In many hate speech cases,

Courts must identify the intent of the speaker's words in order to justify the speech as unprotected under the First Amendment. However, according to Hudson, "The U.S. Supreme Court in Renov. ACLU (1997) noted (albeit in a non-hate speech context) that the Internet is entitled to the highest level of First Amendment protection, akin to the print medium…" (Hudson, 2002). As noted in online hate speech cases like *Elonis v. United States*, cases are still being decided today using the same methods used in hate speech cases from the 1960's. Yet, some unprotected speech categories that Courts use to decide (non-virtual) hate speech cases such as, incitement to imminent lawless action cannot be used in an online hate speech case. According to Hudson:

> "Most online hate speech will not cross into the unprotected category of incitement to imminent lawless action because it will not meet the imminence requirement. A message of hate on the Internet may lead to unlawful action at some indefinite time in the future, but that possibility is not enough to meet the highly speech-protective test in Brandenburg." (Hudson, 2002)

If hate speech on the Internet will not meet certain aspects of the Brandenburg test then it could be time for the Supreme Court to create a universal test that distinguishes between protected and unprotected online hate speech. However, legislators across America have begun creating regulations in certain categories of hate speech that affect minors –cyberbullying. The cyberbullying category of unprotected speech is one of the most important topics in this thesis to discuss, as it has been subject to the most statutes and regulations across the country.

Cyberbullying

Cyberbullying has plagued the online community. With the advent of the digital age, minors are using the Internet more than ever to communicate. The Internet and social media has become a huge part of young people's social lives, which has also introduced situations where kids are bullying other kids online. According to information provided by Dean Laura N. Gasaway's Cyberspace law class at University of North Carolina School of Law:

> "Cyberbullies have two significant advantages over traditional bullies. First, the Internet affords a unique anonymity that is unavailable in face-to-face encounters. This sense of anonymity, although often times false, allows cyberbullying behaviors to escalate quickly because there is less incentive for a bully to stop or to adhere to societal norms. Second, the Internet has increased the bully's reach dramatically." (Gasaway et al, 2010)

Stopcyberbullying.org defines cyberbullying:

> "Cyberbullying is when a child, preteen or teen is tormented, threatened, harassed, humiliated, embarrassed or otherwise targeted by another child, preteen or teen using the Cyberbullies can be classmates, online acquaintances, and even anonymous users, but most often they do know their victims.

Some examples of ways kids bully online are:

- Sending someone mean or threatening emails, instant messages, or text messages.
- Excluding someone from an instant messenger buddy list or blocking their email for no reason.

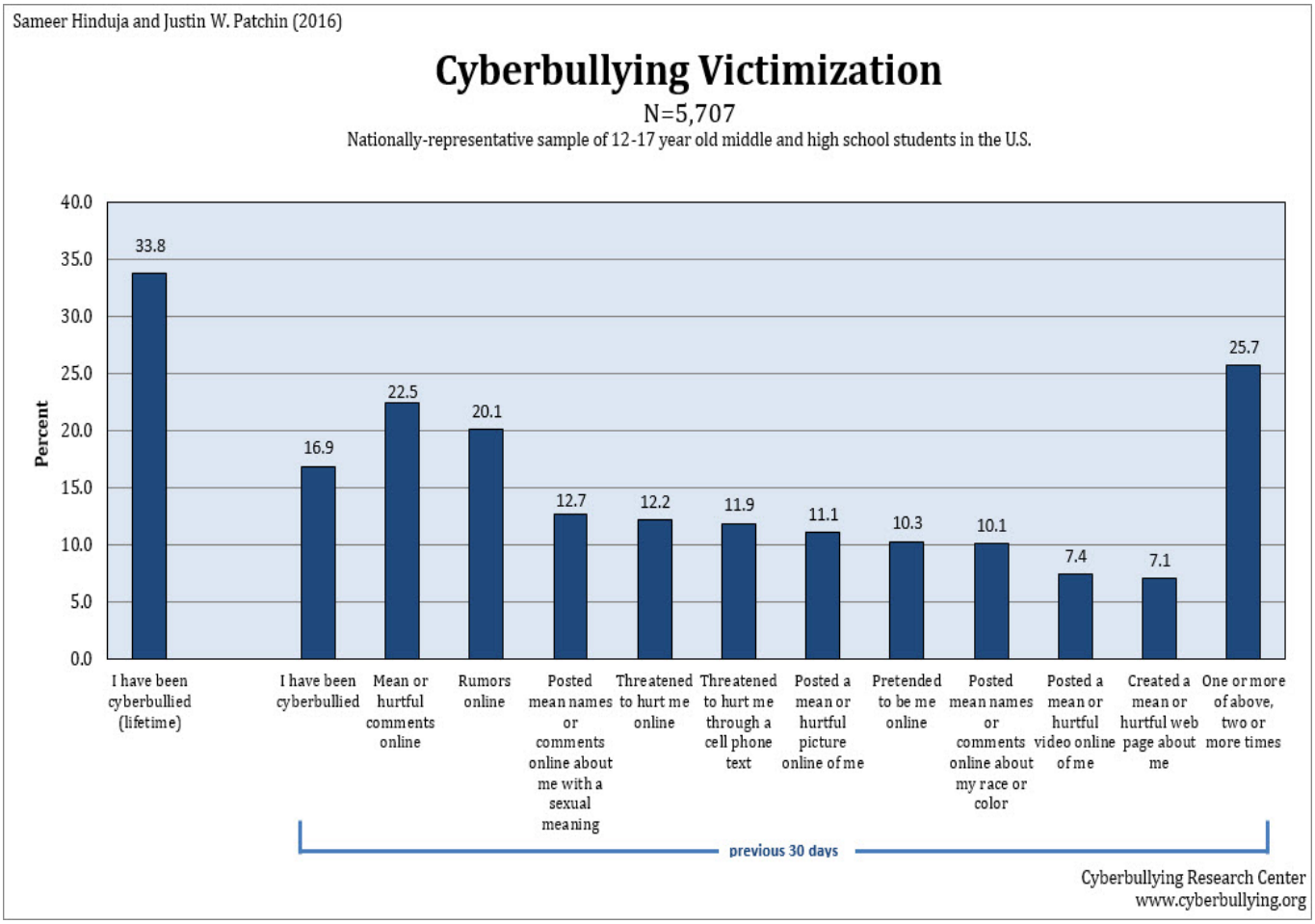
- Tricking someone into revealing personal or embarrassing information and sending it to others.
- Breaking into someone's email or instant message account to send cruel or untrue messages while posing as that person.
- Creating websites to make fun of another person such as a classmate or teacher.
- Using websites to rate peers as prettiest, ugliest, etc.

Cyberbullying is a form of hate speech that is not protected by the First Amendment. Agencies have made recent regulations to protect minors from cyber harassment because, "if schools try to get involved by disciplining the student for cyberbullying actions that took place off-campus and outside of school hours, they are often sued for exceeding their authority and violating the student's free speech right" ("What is cyberbullying, exactly?"). While the First Amendment protects hateful, viscous speech, in recent years U.S. legislatures have passed cyberbullying statutes with a governmental interest to protect minors. According to cyberbullying.org:

> "There is no federal law that specifically applies to bullying. State and local lawmakers have taken action to prevent bullying and protect children. Through laws (in their state education codes and elsewhere) and model policies (that provide guidance to districts and schools), each state addresses bullying differently" ("What is cyberbullying, exactly?").

Cyberbullying statutes are an indication of how bullying and harassment over the Internet truly affects people. According to Sameer Hinduja and Justin W. Patchin from the Cyberbullying Research Center, "Estimates of the number of youth who experience

cyberbullying vary widely (ranging from 10-40% or more), depending on the age of the group studied and how cyberbullying is formally defined…"



"…This study surveyed a nationally-representative sample of 5,700 middle and high school students between the ages of 12 and 17 in the United States. Data were collected between July and October of 2016." (Hinduja and Patchin, 2016).

Although cyberbullying statutes are mainly proposed to protect minors in educational institutions where their speech otherwise would be protected by the First Amendment, these statutes are not universal. After several cyberbullying cases where the victim committed suicide,

states have taken a legislative approach to fighting the issue. According to cyberbullying.org, "50 U.S. states have implemented laws against online harassment since 2011." This can be viewed as conduct vs. speech, where the student's speech was seen as causing harm through its content.

Most of the statutes aimed to update school policies on how to deal with bullying and harassment, rather than criminalizing cyberbullying. Protection against cyberbullying is not a federal law, which, like "true threats" causes irregularity across state Courts in regards to what type of cyberbullying is protected and not protected under the First Amendment. According to cyberbullying.org:

> "All of these [states] require schools to have policies to deal with bullying, and almost all of them refer to electronic forms of harassment (or cyber- bullying specifically), but there exists great variation across states regarding what exactly is mandated…A few states formally criminalize cyberbullying; that is, they specify criminal sanctions such as fines and even jail time for the conduct. In addition, many cyberbullying behaviors already fall under existing criminal (e.g., harassment, stalking, felonious assault, certain acts of hate or bias) or civil (e.g., libel, defamation of character, intentional infliction of emotional distress) legislation, though these laws are infrequently implicated." ("What is cyberbullying, exactly?").

According to Allison Kodjak, a health policy correspondent on NPR's Science Desk, who references an American Academy of Pediatrics report:

> "There is a clear relationship between bullying and suicide thoughts and attempts. Suicidal ideation and behavior were increased in victims and bullies and were highest in

> people who were both bullies and victims of bullying, said the report. And cyberbullying increased suicide attempts as much as face-to-face bullying." (Kodjak, 2016)

While the U.S. Supreme Court has not yet addressed the free speech rights of students in a cyberbullying context, there have been past cases that address student's free speech rights. When it comes to the capability of schools being able to regulate student speech, most cases usually reference *Tinker v. Des Moines Independent Community School District* (1969).

In the case *Tinker v. Des Moines*, student Mary Beth Tinker, who was 13 at the time, successfully challenged a school rule in 1965, which prohibited students from wearing armbands in protest of the Vietnam War. The Supreme Court ruled, "Either students or teachers shed their constitutional rights to freedom of speech or expression at the schoolhouse gate. . . . Students in school as well as out of school are 'persons' under our Constitution [and] are possessed of fundamental rights." (*Tinker v. Des Moines Independent Community School District*, 1969). This case was an influential Supreme Court case because it was the first time the Supreme Court said students have First Amendment rights in schools.

Nevertheless, in this case the Supreme Court also recognized the need "for affirming the comprehensive authority of the States and of school officials, consistent with fundamental constitutional safeguards, to prescribe and control conduct in the schools." In order to balance the opposing interests, the Supreme Court permitted schools to restrict student speech rights if the speech at issue would cause a "material and substantial interference with schoolwork or discipline," or an "invasion of the rights of others." (*Tinker v. Des Moines Independent Community School District*, 1969). With the outcome of this case, it can be argued that if student speech takes place off-campus, and uses no school resources; a school cannot discipline a student

for that speech without violating his/her First Amendment rights. However, Hinduja and Patchin argue that:

> "There are two key features of this case that warrant consideration. First, the behavior considered in Tinker occurred on campus. Second, the behavior was passive and non-threatening. In short, the court ruled that: 'A prohibition against expression of opinion, without any evidence that the rule is necessary to avoid substantial interference with school discipline or the rights of others, is not permissible under the First and Fourteenth Amendments' [emphasis added]. Thus, the Court clarified that school personnel have the burden of demonstrating that the speech or behavior resulted in (or has a reasonable likelihood of resulting in) a substantial interference. This has become the default standard that schools apply when evaluating their ability to discipline students for their misbehavior…" (Hiduju and Patchin, 2008)

There have been advancements in student speech cases since *Tinker v. Des Moines*, but the pressing consequences of cyberbullying brought on by technological advancements has challenged students' First Amendment rights once again. Courts have given leeway to principles when deciding whether a student's speech is considered cyberbullying, using the Tinker Standard – "substantial disruption". However, should Courts and school officials have the right to decide when to take away a student's free speech rights? In the *case J.S. v. Bethlehem Area Sch. Dist*., a student created a personal website, using his own personal laptop, which contained threatening statements made towards students, teachers, and the principal of the school. The school punished the student, and he challenged the ability of the school to penalize him. However, the state court held that the school's punishment did not violate the student's First

Amendment rights even though the speech was done off-campus because (1) the website solicited funds to hire a hit man to kill the student's math teacher, which caused the teacher to take a medical leave of absence. (2) The website was easily accessible at school, and caused students to suffer anxiety. The nature of the speech was threatening, and the speech caused substantial disruption (Tinker Standard). (*J.S. v. Bethlehem Area Sch. Dist.*, 2002)

More recently, in *Kowalski v. Berkeley County Schools*, a student, Kara Kowalski created an online profile, "S.A.S.H", harassing another student. The Myspace group page was apparently an acronym for "Students Against Shay's Herpes," referring to another Berkeley County Schools' student, Shay N. School officials suspended Kara for violating the school's harassment, bullying, and intimidation policy. Kara sued the school for violating her free speech rights and challenged the punishment. The lower court upheld the suspension and the case was appealed to the Fourth U.S. Circuit Court of Appeals, which affirmed the lower court opinion, stating:

> "Kowalski used the Internet to orchestrate a targeted attack on a classmate, and did so in a manner that was sufficiently connected to the school environment as to implicate the School District's recognized authority to discipline speech which 'materially and substantially interfere with the requirements of appropriate discipline in the operation of the school and collide with the rights of others" (Kowalski v. Berkeley County Schools (2011).

In 2014, New York's highest court became the first U.S. court to recognize whether a New York cyberbullying law interfered with students' First Amendment rights. In *People v. Marquan M.*, 2014 the defendant, a 15-year-old high school student, anonymously posted sexual information about fellow classmates on Facebook. He was criminally prosecuted for

"cyberbullying" under a local law enacted by the Albany County Legislature. In 2010, the Albany County Legislature initiated cyberbullying laws to tackle “non-physical bullying behaviors transmitted by electronic means.” The law defined cyberbullying as:

> “Any act of communicating or causing a communication to be sent by mechanical or electronic means, including posting statements on the Internet or through a computer or email network, disseminating embarrassing or sexually explicit photographs; disseminating private, personal, false or sexual information, or sending hate mail, with no legitimate private, personal, or public purpose, with the intent to harass, annoy, threaten, abuse, taunt, intimidate, torment, humiliate, or otherwise inflict significant emotional harm on another person. The law made cyberbullying a misdemeanor offense punishable by up to one year in jail with a $1000 fine.” (*People v. Marquan M.*, 2014)

At trial, the defendant argued that the law violated the Free Speech Clause of the First Amendment because it was overbroad and unlawfully vague. He further explained that the law was overbroad because it prohibited his protected expression. In a 5-2 decision, the New York Court of Appeals agreed with the defendant, concluded that the law was "overbroad and facially invalid under the Free Speech Clause of the First Amendment.

The case, *People v. Marquan M*. demonstrates the difficulty of creating a federal law. Laws that regulate cyberbullying face a delicate balancing act. They must be written narrowly to avoid intruding on speech protected by the First Amendment while still restricting the objectionable conduct. Though, if Courts have generally maintained the rational discipline of students whose cyberbullying was threatening and substantially disrupted the learning

environment at school, and most states have cyberbullying statutes then maybe it is time for congress to pass a universal, federal law to combat cyberbullying.

As of now, cyberbullying is being treated as a misdemeanor crime or civil issue. U.S. Courts are known to support First Amendment rights of free speech; however, there is also a compelling governmental interest to protect children. Under the Free Speech Clause of the First Amendment, the government generally "has no power to restrict expression because of its message, its ideas, its subject matter, or its content" (*United States v. Stevens*, 2010). Yet, the government undeniably has a compelling interest in protecting children from harmful publications or materials, which the court made clear in the case *Reno v. American Civil Liberties Union,*1997). Cyberbullying is not theoretically protected from government regulation, so state legislators must decide for themselves if the First Amendment permits the prohibition of cyberbullying directed at children, depending on how that activity is defined. However, it is clear that just like other categories of unprotected speech, state legislators need to find a way to draw the line between protected speech and cyberbullying on the Internet. Through many of the highlighted cases in this section, it is clear that the line between cyberbullying and protected speech on the Internet is blurred.

<u>Main Point # 3</u>: Does the fact that the Internet and Social Media permit anonymity and a broader audience, enable hate speech? What are the arguments for and against regulating hate speech on social media?

<u>Does the Internet/Social Media enable hate speech?</u>

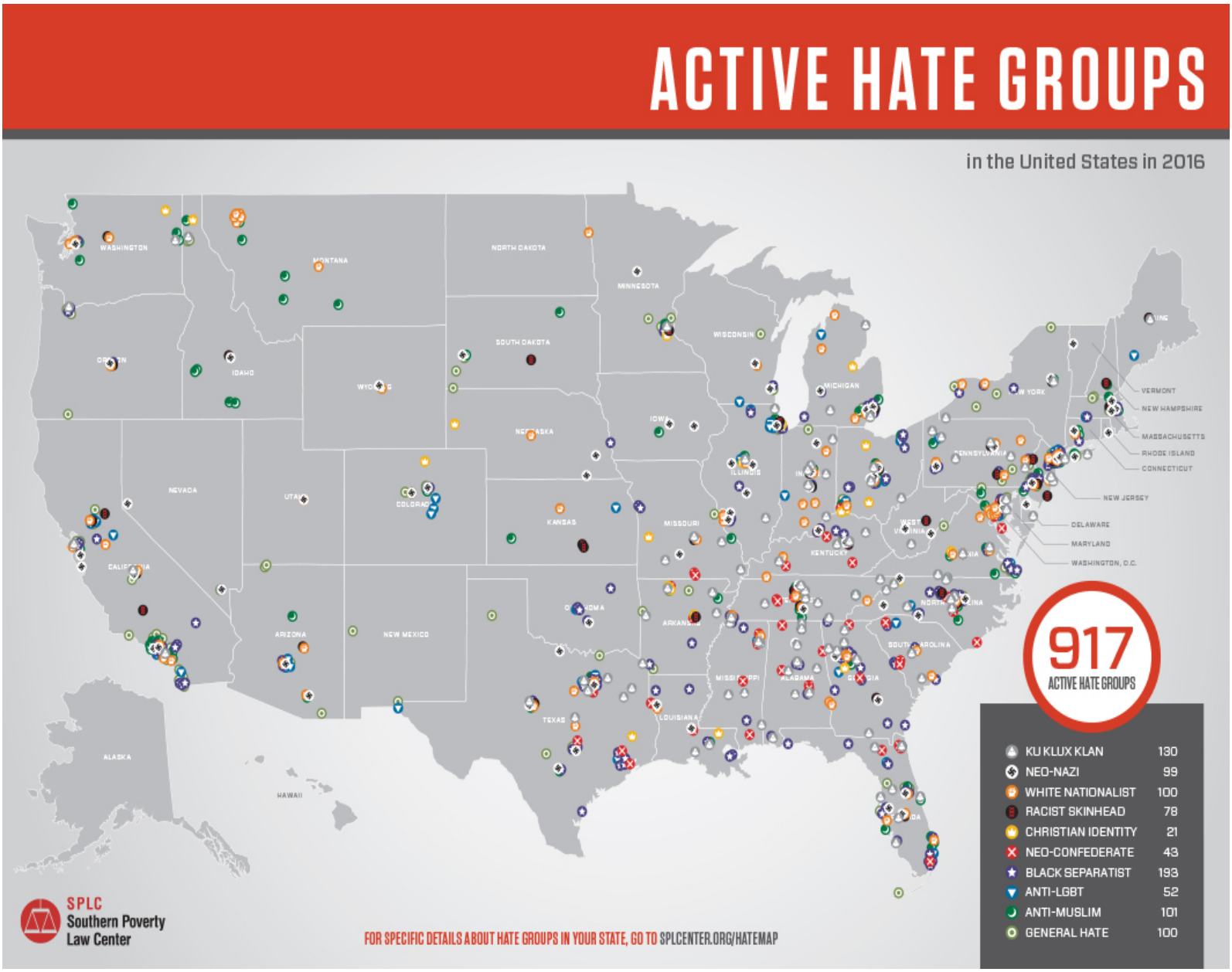
The very definition of hate speech, already vague, became even more complicated when hate

speech made its way to the Internet, along with hate groups. Hate speech on the Internet could be created by anyone, anywhere in the world, posted to a server anywhere around the globe, and be accessed by or targeted at anyone. The Internet allows for a wider audience; therefore, thousands of people could possibly be more susceptible to hate speech than they would have been if they were offline.

The First Amendment states: "Congress shall make no law … abridging the freedom of speech, or of the press," yet it becomes exceedingly complex in its application to Social Media. With the advent of the digital age came the emergence of the personal computers, the Internet, and social media, which are the new media outlets for this current generation. According to a study done by Pew Research Center:

> "Nearly two-thirds of American adults (65%) use social networking sites, up from 7% when Pew Research Center began systematically tracking social media usage in 2005. Young adults (ages 18 to 29) are the most likely to use social media – fully 90% do. Still, usage among those 65 and older has more than tripled since 2010 when 11% used social media. Today, 35% of all those 65 and older report using social media, compared with just 2% in 2005." (Perrin, 2015)

Included in that 65% is a handful of extremists who have taken advantage of the open forums that social media and the Internet provide. According to the Southern Poverty Law Center's active hate group map, which was compiled using hate group publications and websites, citizen and law enforcement reports, field sources and news reports. Logged activity can include criminal acts, marches, rallies, speeches, meetings, leafleting or publishing.



Also according to the Southern Poverty Law Center, a website dedicated to fighting hate and bigotry and to seeking justice for the most vulnerable members in society, there are currently 917 hate groups in America:

> "The SPLC has documented an explosive rise in the number of hate groups since the turn of the century, driven in part by anger over Latino immigration and demographic projections showing that whites will no longer hold majority status in the country by around 2040. The rise accelerated in 2009, the year President Obama took office, but declined after that, in part because large numbers of extremists were moving to the web and away from on-the-ground activities. In the last two years, in part due to a presidential campaign that flirted heavily with extremist ideas, the hate group count has risen again." ("Hate Groups 1999-2016")

Independent Lens Interactive Editor, Craig Phillips; defines a hate group as "an organized group or movement that advocates and practices hatred, hostility, or violence towards members of a race, ethnicity, nation, religion, gender, gender identity, sexual orientation or any other designated sector of society." (Phillips, 2016). No longer do these hate groups have to pass out leaflets on street corners to spread their bigotry, the Internet has created a space for extremists to gain large like-minded audiences across the globe, anonymously able to promote and recruit for their cause- with only the click of a button. According to Phillips:

> "Many would argue that the Internet has been both a blessing and a curse; that it's led to a rise in hate groups and activity, but has also helped expose groups that were already there. As the Internet became a more widely used World Wide Web in the '90s, the first white supremacist hate group to take full advantage was Stormfront, whose website launched in March 1995; after that, there was increasing interest by extremist groups who saw that the Internet could become a major recruiting tool." (Phillips, 2016)

Hate groups play an active role in spreading vitriol online. Yet, hate groups like the Ku Klux Klan's (KKK) conduct and speech – that do not incite violence - are fully protected under the first Amendment offline. So, should it be any different online? Today, hate groups are able to reach a wider audience over the Internet and social media. If a person writes their name on a piece of paper, people know who wrote it. Likewise, if a person gave a speech, people know see who is talking. However, on the Internet anyone can call himself or herself anything they want and can create any website they want – anonymously. Anonymity gives online users freedom from responsibilities, allowing them to say things they otherwise would not dare say in real life. Since users are anonymous, no one will be able to track or trace back actions to their physical selves in real life, and they do not have to face the consequences of their action. As a direct result, the fear of consequence disappears; and the worst side of human nature thrives. In an article written by Alexandar Tsesis, a professor at Loyola University Chicago, School of Law, and the author of "Hate in Cyberspace: Regulating Hate Speech On the Internet," wrote:

> "The relatively inexpensive technologies necessary to run computer servers have enabled hate groups to rapidly increase their presence on the Internet by spreading ideologies through electronic pamphlets, books, and a variety of multimedia documents. They can also engage in real time discussions with similarly minded ideological devotees, even though they are physically hundreds of miles apart. These group meetings, think tanks, and strategy sessions can either be public or the messages can be encrypted for secure conversations with limited audiences" (Tsesis, 2001)

The Internet has clearly played a role in making hate groups more relevant, and making their hate messages more accessible to Internet users. According to Mark Weitzman, the director of the task force against hate and terrorism at the Simon Wiesenthal Center, in an interview with Imaeyen Ibanga, a web and video producer for NBCNews.com; “Stormfront, which encourages children to print out this information and take it to school, is recognized as the first online hate site. Hate groups began using the Web from the very beginning, even before most people had access to the Web in their homes,” Weitzman said. He continued; "You don't have to go hunting for it. It goes right into your house." (Ibanga, 2009)

Although the meaning behind the speech may be similar, online hate speech and spoken hate speech or speech plus conduct differ in content delivery. The Internet allows for anonymity and a wider audience. Anita M. Samuals, freelance journalist and cultural critic, recently wrote a book that explores racism and bigotry in reader comment sections. In *Rants & Retorts: How bigots got a monopoly on commenting about news online*, Samuels explains that there are two types of people who post racist comments online: 1) So-called “regular folks” who would never speak in public about their real views about minorities. 2) White supremacists who “double dip” by going on known hate websites and to news websites. (Samuels 7). According to a quote in the book from Robert S. Anthony, a syndicated columnist who writes the blog “The Paper PC,” “That is no longer the case, of course. Today anti-black comments are often quickly challenged by other comments. While the Internet can be anonymous, its also transparent, and most of todays users seem savvy enough to see through hateful messages.” (Samuels 3).

The book also features an interview with Keith Woods, vice president for diversity in news and operations at National Public Radio in Washington, D.C., who explains how the Internet contributes to the rise in hate speech:

> "The Internet makes it possible to spread all manner of vile ideas about all groups – just as it's made possible to learn the most obscure facts in a matter of seconds…what the Internet provides is 'cover'. People can instantly post anonymously (or using pseudonyms) without having to face consequences. They could always do that with phone calls or mail, but they can do it with virtually no hassle now." (Samuels 9)

Hate groups and extremists have taken over the Internet and social media sites such as Facebook and Twitter, proving that it may be impossible to regulate violent speech on the Internet. When someone vocalizes a hateful or vicious comment, face-to-face, it would normally be a single isolated instance. The government can limit some protected speech by imposing "time, place and manner" restrictions with a governmental interest in protecting citizens. However, it is much more difficult to apply those same restrictions on the Internet because "it allows for a consistent course of conduct rather than a single isolated instance" (*Time, Place, and Manner Restrictions*). The government may never be able to impose a time, place, and manner restriction on the Internet. According to onlinerights.org: "Time, place, and manner restrictions are typically context-specific and relate to the needs of a particular area, venue, or community. Because of the global nature of the Internet, it would be very difficult to develop or enforce meaningful time, place, and manner restrictions online" (*Time, Place, and Manner Restrictions*).

Besides time, place and manner restrictions, Courts struggle with enforcing laws against online harassment for a variety of reasons. According to partnersagainsthate.org, a collaborative project of the Anti-Defamation League, The Leadership Conference Education Fund, and the Center for Preventing Hate:

> "A difficulty in enforcing laws against harassment is the ease of anonymous communication on the Internet. Using a service that provides almost complete anonymity, a bigot may repeatedly e-mail his or her victim without being readily identified." (Partners Against Hate)

The Internet offers opportunities for anyone who wants to express an opinion about anything to make their thoughts accessible to a world-wide audience far more easily than before the digital age. Through research provided in this thesis, it has become very clear that anonymity, the global nature of the Internet, and accessibility have played a huge role in the growth and spread of hate on social media sites.

Now that we have discussed if the Internet and Social Media enable online hate speech, let's look at the arguments for and against regulation hate speech online.

Arguments for and against regulating hate speech online

Dissemination of hate speech has a long history in this country, with many Americans advocating the right to free speech, including hateful speech. Although the platform for the speech may have changed through the decades, much of the feeling has remained the same.

The online hate speech epidemic has led to intense debates across the country about the need for some form of regulation. The controversy surrounding hate speech has led lawmakers to reexamine some of the most fundamental principles on which America was built. When dealing with hate speech cases, especially online, Courts must find a way to balance freedom of expression and respect for basic human rights such as fairness and dignity.

So then how does Congress find a balance that protects citizen's civil rights without limiting the civil freedoms of the speaker? It seems as though that balance may not be possible as the judicial system, lawmakers, and ordinary citizens have been questioning the need to regulate the growing number of threats being made online. With movements like the Men's Rights movement and White Lives Matter becoming increasingly popular, minorities are constantly being targeted and threatened online. People are questioning how to address the threats before they become part of the norm. Yet, how do we address these issues when the laws were written before the Internet existed and Courts have not figured out a way to tackle threats made online?

Some Americans believe all speech should be protected under the First Amendment, no matter how disgusting or hateful it is, while others believe laws regulating online hate speech are long overdue. The debate over hate speech have been going on even before the arrival of the digital age. According to M. Alison Kibler, who has been referenced earlier in this thesis:

> "History clearly shows, however, that struggles over hate speech are nothing new. Hate speech has been a century-long rift in American politics because it pits two deeply held American values against each other: free speech and equality." (Kibler, 2015)

Although American society has evolved since the founding era, the debate over the regulation of hate speech today remains the same. The American Civil Liberties Union, an organization that works to defend and preserve the individual rights and liberties guaranteed by the Constitution and laws of the United States states that:

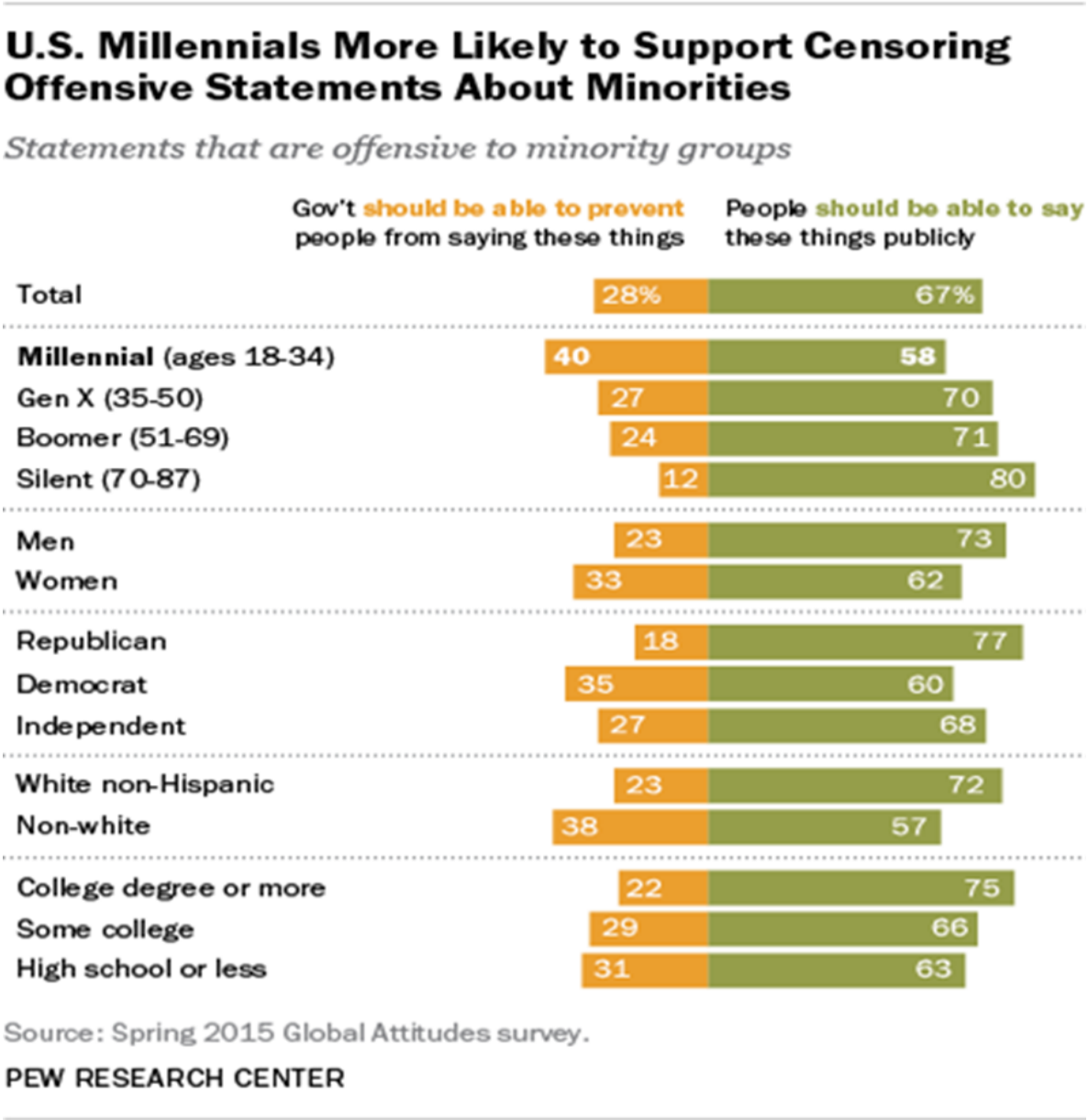
> "Free speech rights are indivisible. Restricting the speech of one group or individual jeopardizes everyone's rights because the same laws or regulations used to silence bigots can be used to silence you. Conversely, laws that defend free speech for bigots can be used to defend the rights of civil rights workers, anti-war protesters, lesbian and gay activists and others fighting for justice." (American Civil Liberties Union*, Hate Speech on Campus)*

Many Americans seem to agree with the ACLU. According to a 2015 Pew Research Center survey:

> "Americans were among the most supportive of free speech, freedom of the press and the right to use the Internet without government censorship. Moreover, Americans are much more tolerant of *offensive* speech than people in other nations [are]. For instance, 77% in the U.S. support the right of others to make statements that are offensive to their own religious beliefs, the highest percentage among the nations in the study. Fully 67% think people should be allowed to make public statements that are offensive to minority groups, again the highest percentage in the poll." (Wike, 2016)

However, in another survey conducted by the Pew Research Center, 40% of millennials (ages 18-34) want the government to step in and censor speech that is offensive to minority groups on the Internet. According to the survey: "Four-in-ten Millennials say the government should be able to prevent people publicly making statements that are offensive to minority groups, while 58% said such speech is Ok. Even though a larger share of Millennials favor allowing offensive speech against minorities, the 40% who oppose it is striking given that only

around a quarter of Gen Xers (27%) and Boomers (24%) and roughly one-in-ten Silents (12%) say the government should be able to prevent such speech." (Poushter, 2015)



Although the public's response to handling hate speech on the Internet differs, we must focus on how online threats are currently being handled. With no clear definition of what constitutes a true threat, or clarification between protected and unprotected speech on the Internet - which has been proved with various cases in this thesis, one question still remains: Do we need

different rules for online threats? Many can agree that most threats made online are just trolls (defined by the Oxford Dictionary as: a person who makes a deliberately offensive or provocative online post) being trolls. Why not just leave the laws that regulate unprotected speech the way they are? Why not let Courts and the police determine which cases of threatening speech might escalate into offline behavior that could potentially cause someone harm? In other words, how can the court decide online hate speech cases if there is no definition of what constitutes hate or any federal/universal tests and regulations that Courts can use when decided these cases?

It is common to call the police when someone is in a threatening situation. However, cops may ignore online threats because there is no perceived imminent threat, regardless of how violent it may be. It is not until someone actually commits a hate crime that our government decides to step in and figure out where the plaintiff's ideas may have come from. If someone is being threatened on a social media site, the answer could be as simple as avoiding the site that they are being harassed on. However, do innocent people deserve to be driven off a popular, public forum in fear for their life? What happens when hate speech does more than simply express differing views and starts to promote fear and intimidation? This may result in emotional harm of hate speech victims.

There is no clear definition of what exactly constitutes harm on the Internet. Courts have found a small category of speech that is not protected by the First Amendment, such as speech that is considered a true threat, cyberbullying, or intent to harm. However, what exactly constitutes harm?

In other words, does harm mean physical harm, or does it include emotional harm? Should we consider it harmful that the threats are viewable to anyone online, instead of just

telling targets to avoid them? Harassment and violent speech over the Internet causes stress and fear for the victims. Is that not considered harm? Through studies included in this thesis, we see that women and people of color, are often the ones who are the target of online hate speech. Yet, minorities' lives and experiences are marginalized in American society. Is that not considered harm, to continue to berate and degrade a group of people who are constantly disregarded in the real world? It is also potentially harmful to post people's personal information like addresses and phone numbers on public sites; however people do it regardless, and with little repercussion. Does that not put a person in potential danger, even if the harm that can come from it is not imminent? Effectively, the question then becomes, how do we weigh the emotional harm of the victim of hate speech against the defendant's First Amendment right to speak freely?

The 14th Amendment states:

> "No state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United State; nor shall any state deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws." (Elbel, 2008)

The equal protection clause of the 14th Amendment prohibits discrimination, for it provides individuals or classes with "equal application" of the laws. Some argue that hate speech infringes the 14th Amendment's guarantee of equal protection under the law. According to Eugene Volokh, a Gary T. Schwartz Professor of Law at UCLA School of Law:

> "Hate speech infringes the Equal Protection Clause because individuals have a right to be protected against bigoted and pornographic speech. This right derives from the concerns that: (1) when a bigoted society is partially constructed due to unlimited bigoted speech, individual rights to equal protection are interfered with, and (2) a victim's own free speech rights are interfered with when bigoted speech causes individuals to fall silent or devalues their voices in the marketplace of ideas." (Volokh, 1996)

There is no easy way to decide whether hate speech should be regulated or if it should continue to be protected by the First Amendment. One of the main problems with hate speech laws is defining what constitutes hate - it is highly subjective. However, Courts have failed to draw the line between protected and unprotected online speech with certainty, clarity and consistency. Although the U.S. Supreme Court made their decision in the case *Elonis vs. United States* that hate speech is protected under the First Amendment, this author proposes that Congress needs to create a universally accepted definition of hate speech, thus making it easier to propose federal laws and create new tests/standards to guide lower Courts. The term harm also needs to be appropriately defined, because online speech can both emotionally and mentally harm users. Free speech is not an absolute right, and if that is the case then Congress has a duty to the American people to define hate speech and harm.

<u>Main Point # 4)</u> What implications are there, if any, for social media platforms such as Twitter and Facebook when it comes to dealing with hate speech?

Not only has there been a rise of hate sites on the Internet, there has also been an increase in hateful insults received through social media. Social media facilitates the rapid spread of hate

speech on the Internet, and platforms like Facebook and Twitter have become prime new ground for hate speech – with more users joining and creating profiles daily. Social media platforms have become a cesspool of vile vitriol, and according to Huffington Post journalist Sean McElwee:

> **"** Humboldt State University compiled a visual map that charts 150,000 hateful insults aggregated over the course of 11 months in the U.S. by pairing Google's Maps API with a series of the most homophobic, racist and otherwise prejudiced tweets." (McElwee, 2013)

The data collected from the map proves that hate speech is not going to vanish from social media on its own. Social media platforms have found it necessary to implement policies that will help identify hate speech and abuse. Twitter has created its Trust and Safety Counsel in order to create a safer place for individuals to express themselves without causing harm to others. Social Media plays an essential role in today's society by allowing the public to display their own opinions, which begs to question; does social media have the right to censor their users' content?

Social media sites such as Facebook and Twitter may voluntarily agree to prohibit users from sending racist or bigoted messages over their service because the government plays no role in its actions. Twitter and Facebook have no First Amendment obligation to let users use their services and will censor speech whenever they choose to. Social media sites get to make their own set of speech rules and among those rules are prohibitions on using their services to harass or spread vitriol.

Hateful comments and Twitter trolls are nothing new to Twitter; however, the 2016 presidential election brought online hate speech more into the spotlight due to Trumps infamous tweets. According to a report published by the Anti-Defamation League, “during the U.S. presidential election, from August 1, 2015 through July 31, 2016, there were over 2.6 million tweets "containing language frequently found in anti-Semitic speech. Additionally, at least 800 journalists received anti-Semitic tweets. There is evidence that a considerable number of the anti-Semitic tweets targeting journalists originate from people identifying themselves as Trump supporters, “conservatives” or extreme right-wing elements.” (*Anti-Semitic Targeting of Journalists during the 2016 Presidential Campaign,* 1)

Not only have journalists and minorities become targets of hate speech on Twitter; hate speech is being directed at celebrities, many of whom use social media to connect with their fans and promote themselves and their work. Leslie Jones, a cast member on "Saturday Night Live" and star of the "Ghostbusters" remake, was driven off Twitter in July, 2016 after receiving racist and sexist tweets – just for being an African-American female. Jones signed off from Twitter. "I leave Twitter tonight with tears and a very sad heart. All this cause [sic] I did a movie. You can hate the movie but the shit I got today…wrong," she tweeted. According to a New York Times article written by Katie Rodgers:

"Several users compared her to primates, including Harambe, the male gorilla who was shot dead in May at a Cincinnati zoo. Others sent her pornography. Another created a fake Twitter account that appeared to show the actress spewing hateful language…Twitter permanently banned [Milo]Yiannopoulus [who heavily harassed Jones] from its platform. On Tuesday, Twitter said in a statement: 'This type of abusive behavior is not permitted on Twitter, and we've taken action on many of the accounts reported to us by both Leslie and others…ranging from warnings that also require the deletion of Tweets violating our policies to permanent suspension." (Rodgers, 2016)

Over the past year, Twitter has been taking harassment very seriously - it has banned revenge porn, issued new anti-harassment rules, established a trust and safety council and deleted the accounts of high-profile users like Milo Yiannopoulos, who it considers abusive. As explained in the Twitter Rules under the hateful, conduct section:

"You 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 disease. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories. Examples of what we do not tolerate includes, but is not limited to behavior that harasses individuals or groups of people with:

- Violent threats
- Wishes for the physical harm, death, or disease of individuals or groups;
- References to mass murder, violent events, or specific means of violence in which/with which such groups have been the primary targets or victims;
- Behavior that incites fear about a protected group;
- Repeated and/or or non-consensual slurs, epithets, racist and sexist tropes, or other content that degrades someone." (*Twitter Terms of Services*, 2016)

Facebook has similar terms of services:

"We work hard to make our platform a safe and respectful place for sharing and connection. This requires us to make difficult decisions and balance concerns about free expression and community respect. While there is no universally accepted definition of hate speech, as a platform we define the term to mean direct and serious attacks on any protected category of people based on their race, ethnicity, national origin, religion, sex, gender, sexual orientation, disability or disease…We remove credible threats of physical harm to individuals…We don't tolerate bullying or harassment. We allow you to speak freely on matters and people of public interest, but remove content that appears to

> purposefully target private individuals with the intention of degrading or shaming them. (*Facebook Terms of Service*, 2016)

Social media companies are under scrutiny by users and human rights activists for two reasons: 1) not doing enough to combat the amount of hate speech that thrives on their platforms, and 2) restricting user's rights to free speech. Companies now have to deal with a difficult balance between protecting freedom of expression for users while also creating an open and welcoming community. Celebrities like Leslie Jones, and minorities who are constantly being harassed on their social media sites, are calling for stricter regulation of Internet speech on social media platforms. However, users who are also harassed on the site offer other solutions rather than the company suppressing hateful speech. In an article written by New York Times reporter Jim Rutenburg, and Jeffrey Goldberg, a national correspondent for The Atlantic, makes an interesting point:

> "Mr. Goldberg says he is torn about what Twitter should do, given that its cause — openness and free speech — is a reason he and so many other journalists are drawn to the service. "That's the fundamental problem," he told me. "At a certain point I'd rather take myself off the platform where the speech has become so offensive than advocate for the suppression of that speech." (Rutenburg, 2016)

That could be a simple solution, but should users like Leslie Jones be driven off their social media sites for being the victims of harassment. There is a huge difference between "I didn't like that movie" and "This person should die because of their race." Although the

Supreme Court and Congress has not been able to draw the line between unprotected and protected speech online, social media sites have begun to on their platforms. Facebook and Twitter are just a few social media sites that are taking a stance against online hate speech. They are punishing users for their actions by deleting accounts, or removing the speech that violates their policies. However, social media sites may be inadvertently aiding the divide in America by choosing to take away user's freedom of speech. According to Rutenburg;

> "It's all a bit tricky for a company founded with an absolutist ethos, once calling itself 'the free speech wing of the free speech party.' Some of its moves to curtail abuse have drawn accusations that it is applying a double standard aimed at conservatives…Banned Twitter provocateur, Charles C. Johnson...says he is planning a lawsuit to fight his suspension. In an interview, Mr. Johnson said he respected Twitter's right to ban patently offensive speech but argued that it needed to set a consistent, uniformly applied standard. Still, he said, "the problem of trolls" might be unsolvable." (Rutenburg, 2016)

Although a hard task, social media sites continue to step into the uncharted territory of restricting hate speech on their platforms. However, social media sites make their own rules, and are currently combating online hate speech in ways the U.S. government simply cannot because the First Amendment only applies to the government, not private companies. They have created clear guidelines of what kinds of speech are permitted on their sites, and have come up with their own definition of hate speech. However, as Charles C. Johnson points out, the policies are often inconsistently applied. According to Caitlin Elizabeth Ring, a former University of Boulder PhD student who wrote her dissertation on hate speech in social media:

> "Many of the decisions made by the content removal teams at these organizations are not nearly as speech protective as the First Amendment and U.S. Supreme Court precedent on the subject would mandate. Thus, the current situation gives social media companies' unprecedented power to control what videos, text, images, etc. users may or may not post or access on those social media sites." (Ring, 2013)

A national survey of 1,520 adults conducted between March 7-April 4, 2016, found that nearly eight-in-ten online Americans (79%) now use Facebook, more than double the share that uses Twitter (24%) (Greenwood et al, 2016). So, what do social media companies' policies and regulations mean for American users? It's up in the air at this point. However, if Congress created a universal definition of hate speech and clarified the difference between unprotected and protected speech, then maybe dealing with hate speech on the Internet would be less confusing for all.

<u>Main Point #5:</u> Is there a solution that will minimize hate speech on social media, while finding ways to not invade individuals' First Amendment rights?

Since the First Amendment allows the right to free speech, it places limits on laws that attempt to control online speech. So, is there an effective way to deal with online hate speech without invading American's First Amendment rights? If we look at it from a legal standpoint, the solution could be to create strategic laws. The article, "Debating the Mighty Constitutional Opposites", suggests legal options:

> "One way to deal effectively with hate speech is to create laws and policies that discourage bad behavior but do not punish bad beliefs. Another way of saying this is to create laws and policies that do not attempt to define hate speech as hate crimes, or "acts." In two recent hate crime cases, the U.S. Supreme Court concluded that acts, but not speech, may be regulated by law." (*Debating the Mighty Constitutional Opposites*)

Others believe there are additional solutions to counter online hate speech besides criminalizing it. In an article written by David L. Hudson Jr, First Amendment scholar, he includes Christopher Wolf, chair of the Anti-Defamation League's Internet Task Force opinion on how to regulate online hate speech:

> "There is a wide range of things to be done, consistent with the First Amendment, including shining the light on hate and exposing the lies underlying hate and teaching tolerance and diversity to young people and future generations," he said. "Counter-speech is a potent weapon." (Hudson, 2002)

European countries do not have the same freedoms that the U.S Constitution affords its citizens, which makes it a lot easier for their governments to pass laws prohibiting online hate speech. However, the Council of Europe's Additional Protocol to the Convention on Cybercrime has at least clarified a definition for online hate speech, something that the American government has yet to clarify. It defines online hate speech as:

> "any written material, any image or any other representation of ideas or theories, which advocates, promotes or incites hatred, discrimination or violence, against any individual

or group of individuals, based on race, colour, descent or national or ethnic origin, as well as religion if used as a pretext for any of these factors." (Council of Europe, 1997)

Conclusion

If the American government took the same steps to clarify and define online hate speech, then dealing with online hate speech would be less confusing for all. After examining different cases and legislations involving online hate speech, this author has arrived at the position that the American government needs to clarify when free speech crosses over to non-protected speech that is considered a "true threat" on the Internet. However, I propose that the government needs to start by defining it. Without a national definition there continues to be inconsistencies in the law regarding what is considered protected speech on the Internet. I also propose that there needs to be guidelines and direction from the Supreme Court to assist and guide lower Courts when deciding on hate speech cases. All Courts, state legislators, Congress and Internet users should know when free speech crosses over to non-protected speech that is considered a "true threat" on the Internet.

Works Cited

Anti-semitic Targeting of Journalists During the 2016 Presidential Campaign: A Report from ADL's Task Force on Harassment and Journalism. Anti-Defamation League, 2016.

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

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

Chemaly, Soraya , and Mary Anne Franks. "Supreme Court May Have Made Online Abuse Easier." Time, 3 June 2015. Web. 19 Apr. 2017.

Colb, Sherry F. "The Supreme Court Considers “True Threats” and the First Amendment." Verdict Comments. Justia, 06 Dec. 2014. Web. 19 Apr. 2017.

Council of Europe. Committee of Ministers. Recommendation 97(20) on “hate speech”. October 20, 1997.

"Debating the “Mighty Constitutional Opposites”." ABA Division for Public Education: Students: Debating the "Mighty Constitutional Opposites": Hate Speech Debate. ABA Divison for Public Education, n.d. Web. 30 Oct. 2016.

Dennis v. United States, 341 U.S. 494 (1951)

"Determining the Legality of a threat." Numberg Files. Standford.edu, n.d. Web. 19 Apr. 2017.

Dussich, Rebecca . "True Threats and the First Amendment: Objective vs. Subjective Standards of Intent to Be Revisited in Elonis v. United States." The UC Law Review Forum. UCLA Law Review, 05 Jan. 2015. Web. 19 Apr. 2017.

Elbel, Fred. "The US Constitution: 14th Amendment." The 14th Amendment. N.p., 1 Dec. 2008. Web. 19 Apr. 2017.

Elonis v. United States, 575 U.S. ___ (2015).

"Elonis v. United States." Harvard Law Review Elonis v United States Comments. N.p., 10 Nov. 2015. Web. 05 Nov. 2016.

Eugene Volokh, Freedom of Speech and the Constitutional Tension Method, 3 U. CHI. ROUNDTABLE 223 (1996)

"Facebook." Terms of Service. N.p., n.d. Web. 05 Nov. 2016.

Fisher, William. "Freedom of Expression on the Internet." Freedom of Expression on the Internet. N.p., 14 June 2001. Web. 30 Oct. 2016.

Gasaway, Laura N. "Cyberbullying." Cyberbullying - Home. UNC School of Law , May 2010. Web. 19 Apr. 2017.

"Hate Groups 1999-2016." Southern Poverty Law Center. SPLC, n.d. Web. 19 Apr. 2017.

"Hate Speech on Campus." American Civil Liberties Union. N.p., n.d. Web. 30 Oct. 2016.

Hinduja, S. & Patchin, J. W. (2008). Cyberbullying fact sheet: A brief review of relevant legal and policy issues. Cyberbullying Research Center. Web. 19 Apr. 2017.

Hinduja, S. & Patchin, J. W. (2014). Cyberbullying fact sheet: Identification, Prevention, and Response. Cyberbullying Research Center. Web. 19 Apr. 2017.

"How do Courts determine whether speech is a true threat?" First Amendment Schools, n.d. Web. 19 Apr. 2017.

Hudson, David L., Jr. "Hate Speech Online." FirstAmedmentcenter.org. First Amendment Center, 13 Sept. 2002. Web. 30 Oct. 2016.

Hudson, David L., Jr. "True Threats." NewseumInstitute.org. Newseum Institute, 12 May 2008. Web. 19 Apr. 2017.

Ibanga, Imaeyen. "Hate Groups Effectively Use Web as a Recruiting Tool." ABC News. ABC News Network, 12 June 2009. Web. 19 Apr. 2017.

Is there any kind of hate speech on the Internet that is not protected by the First Amendment? Partners Against Hate, n.d. Web. 19 Apr. 2017.

J.S. v. Bethlehem Area Sch. Dist., 807 A.2d 847 (Pa. 2002).

Kibler, Alison. "The Long History of Hate Speech." History News Network. N.p., 19 Mar. 2015. Web. 30 Oct. 2016.

Kodjak, Alison. "Teen Bullies And Their Victims Both Face A Higher Risk Of Suicide." Health News From NPR. NPR, 28 June 2016. Web. 19 Apr. 2017.

Kowalski v. Berkeley County Schools, 652 F. 3d 565 - Court of Appeals, 4th Circuit (2011).

Liptak, Adam. "Supreme Court Overturns Conviction in Online Threats Case, Citing Intent." The New York Times. The New York Times, 01 June 2015. Web. 19 Apr. 2017.

McElwee, Sean. "The Case for Censoring Hate Speech." The Huffington Post. TheHuffingtonPost.com, 24 July 2013. Web. 30 Oct. 2016.

People v. Marquan M., 2014 WL 2931482 (Ct. App. NY July 1, 2014).

Perrin, Andrew. "Social Media Usage: 2005-2015." Pew Research Center: Internet, Science & Tech. Pew Research Center, 08 Oct. 2015. Web. 19 Apr. 2017.

Phillips, Craig. "Tracking Hate Groups Online." Independent Lens. Public Broadcasting Service, 1 Apr. 2016. Web. 19 Apr. 2017.

Poushter, Jacob. "40% of Millennials OK with limiting speech offensive to minorities." Pew Research Center. N.p., 20 Nov. 2015. Web. 19 Apr. 2017.

Reno v. American Civil Liberties Union, 521 U.S. 844, 875 (1997).

Ring, Caitlin Elizabeth, "Hate Speech in Social Media: An Exploration of the Problem and Its Proposed Solutions" (2013). Journalism & Mass Communication Graduate Theses & Dissertations. 15.

Rutenburg, Jim. "On Twitter, Hate Speech Bounded Only by a Character Limit." The New York Times. TheNewYorkTimes.com, 2 Oct. 2016. Web. 5 Nov. 2016.

Samuels, Anita M. Rants & Retorts: How bigots got a monopoly on commenting about news online. CreateSpace Independent Publishing Platform, November 6, 2016.

Schenck V. United States, 249 U.S. 47 (1919)

"The Twitter Rules." Twitter Help Center. N.p., n.d. Web. 05 Nov. 2016.

Tinker v. Des Moines Independent Community School District, 393 U.S. 503 (1969).

United States v. Stevens, 559 U.S. 460, 468 (2010).

"Time, Place, and Manner Restrictions." Online Art Rights — A Resource on Art and Free Speech on the Internet. N.p., n.d. Web. 19 Apr. 2017.

Tsesis, Alexander, Hate in Cyberspace: Regulating Hate Speech on the Internet. San Diego Law Review, Summer 2001. Web. 19 Apr. 2017.

Waldron, Jeremy. The Harm in Hate Speech. Cambridge, Mass: Harvard University Press, 2012. Print.

Watts v. United States, 394 U.S. 705 (1969)

"What is cyberbullying, exactly?" STOP cyberbullying. Cyberbullying.org, n.d. Web. 19 Apr. 2017.

Wike, Richard. "Americans more tolerant of offensive speech than others in the world." Pew Research Center. N.p., 12 Oct. 2016. Web. 19 Apr. 2017.