What Policy would be the Most Effective in Mitigating Online Extremism?

Amber Lynn Myles
California Polytechnic University, Pomona

Online extremism is a progressing global threat which requires the attention of policy makers to mitigate its harmful effects on society. With the capability to instantly access media outlets, both terrorists and extremists have the power to spread dangerous content at a more rapid pace than ever before. This thesis explores the three core policies that are debated among governments – censorship, counter-narrative programs, and media accountability. Through data analysis, the successes and limitations of each policy are discussed. This thesis concludes that, amending media protection legislation to include extremist material as unprotected content is the most appropriate policy to implement. Media accountability – as opposed to other policy options – is more practically feasible, constitutionally sound, and will most likely lead to greater security on the general Internet.

Introduction

Online extremism is a recent threat to online security which the government is working to implement an effective policy to combat its effects on society. Since the spark of the digital era in the mid-1990's, new media has become readily available to much of the developed, global population. Though considerable benefits have come from the expansion of the Internet, there have also been many concerning advancements made by extremists and terrorists, as media platforms allow individuals to collect and produce content instantaneously. The National Consortium for the Study of Terrorism and Responses to Terrorism (START) found that social media significantly increases the efficiency of terrorists in their radicalization efforts. In 2005, when social media was just beginning to develop, the radicalization process for extremists living in the United States took approximately 18 months whereas, in 2016, it only took about 13 months for extremists to reach a state of total radicalization (Jensen et al. 2018). Additionally, START recorded that, in 2016, social media played a primary role in the radicalization process of approximately 86.75% of extremists; in 2012, however, only 48% of extremists claimed social media to be a leading source of influence in their radicalization (Jensen et al. 2018). Though Europol (2019) indicated a decline in terrorist activity on the Web throughout Europe in 2019, the overall trend has shown that, as social media outlets expand and gain more users, terrorist and extremist activity also increases significantly. Since Internet platforms face no international barriers, radicals have the power to reach a global audience and are becoming increasingly reliant on such platforms to further their agendas and gain more recruits within their organizations.

Social media carries influence with all forms of extremism, making the issue detrimental to international security since extreme content on the Web is often used as inspiration for real life terrorist attacks. For example, the New Zealand mosque shooting of 2019 was live streamed on Facebook, gained thousands of views, and was reposted hundreds of times before the footage was removed over a half-hour after being streamed (Flynn, 2019). The video undeniably played a role in inspiring several other extremists to conduct their own shootings – specifically one of the more recent instances in El Paso, Texas. Michael Hayden from the Southern Poverty Law Center (SPLC) discussed the similarities between the two shootings based on the shooters' manifestos that were published online. He wrote, "In both attacks, the suspects published manifestos…both manifestos were saturated with white nationalist talking points, portraying whites as the victims of a plan for elimination" (Hayden, 2019b). In another article from SPLC, it was recorded that the New Zealand shooter had been heavily inspired by the Norwegian terrorist Anders Breivik, as well as Dylan Roof, who conducted a church shooting in Charleston, North Carolina in 2015 (SPLC, 2019). The New Zealand shooter had praised both terrorists in his manifesto and engraved white supremacist slogans on his weapons along with names of previous individuals who had conducted shootings against both Jews and Muslims.

Created by A. Lynn Myles, Department of Political Science, California Polytechnic University, Pomona for their senior thesis project. Correspondence in regards to this research paper should be addressed to A. Lynn Myles Department of Political Science, California Polytechnic University, Pomona. Email: almyles@cpp.edu

These patterns of copycat attacks can be argued as one of the most dangerous effects of social media usage by extremists. Oftentimes, terrorists will use similar language in their manifestos or propaganda videos to reiterate their extreme message and gain legitimacy or sympathy among their audiences. Terrorists have also used social media as a recruitment tool for their organizations. The Islamic State (ISIS) is a terrorist organization which has an unmatched fluency in social media (among other terrorist organizations) and has used several new media platforms to its advantage. Similar to right-wing extremists' use of rhetoric that evokes a sense of fear among the White population, ISIS recruiters use a tone among a young audience that will elicit a sense of belonging and brotherhood while also pushing the agenda of fighting for the expansion of ISIS. The recently deceased leader of ISIS, Abu Bakr Al-Baghdadi, claimed that he used social media platforms to appeal to young men and teenagers to join the fight with ISIS; he believed he had a greater opportunity to persuade a global audience through the media rather than by face-to-face encounters (Awan, 2017). Islamic terrorists tend to use common phrases such as "brotherhood" and "victory" in their propaganda to gain enthusiasts among a population with self-identity vulnerabilities (Awan, 2017). NBC news reported that approximately 250 young Americans had tried to join ISIS in 2015, and each suspect had a long history of communication with terrorist recruiters through social media (Engel et al. 2016). Many of those attempting to join ISIS were caught by the FBI; however, with the increasing capabilities of social media platforms, there may be a larger wave of recruits in the future that will be successful in their attempts to join ISIS.

While online extremism is a relatively recent concern for the government, scholars have presented differing theories on how to best address the consequences of new media extremism. Nonetheless, with each policy presented, there are several legal, ethical, and practical challenges that make the issue complex in nature and difficult to develop a single conclusion to address the problem. By understanding what online extremism is, how the Internet is used by terrorists and extremists, and whom the extremism affects, policy makers will have better understanding on how to implement effective legislation. Some scholars believe that censorship is the best method to prevent radicalism, while others argue that there are less controversial ways to mitigate online extremism. There are also some who believe that counter-narrative programs would be the best method in combating extremism and preventing future terrorists from radicalizing, though others argue that this policy is too complex to empirically measure its success. Media accountability is another policy that has been gaining the attention of various governments, but there are legal and practical complexities that would need to be addressed before implementing such a legislation. By examining each policy option in length, the goal of this thesis is to develop a general recommendation as to which policy would be the most effective in mitigating online extremism.

II. Literature Review

The purpose of this literature review is to discuss the debate among scholars regarding the various policy options that are considered in the efforts to reduce online extremism. First, this literature review will summarize who terrorists target and how various extremist groups utilize new media to further their agendas. Second, this literature review will discuss the differing benefits and consequences of censorship, counter-narrative programs, and media accountability debated among scholars. Finally, this review will conclude with the major points that scholars have regarding each of the policy options.

Online Extremism – Overview

The majority of researchers agree that new media grants terrorists and extremists the ability to produce dangerous content at a quicker rate than ever before. Amble's article analyzed jihadists' use of new media since the mid-1990's (Amble, 2012). He claimed that improved digital technology, global expansion of the Internet, and reduced computing prices contribute to providing jihadists with the ideal platform to use the Internet for communication, fundraising, training, and propaganda purposes (Amble, 2012). Nele Schils and Antoinette Verhage (2017) claimed that the Internet provides the public with easy access to network with extremist groups. They used survey data on individuals from ages 16-24 years old who had radical convictions, then held interviews with three groups that were divided into left-wing, right-wing and religious categories. The interviews had some limitations: the short period of time that they were conducted in, a hesitation in participants due to anonymity concerns, and the differing levels of extremism that participants experienced (Schils; Verhage, 2017). However, despite the limitations, Schils and Verhage found consistency in all three groups. They discovered that young individuals in each category were likely to be targeted by extremists due to feelings of injustice towards the government. Various terrorists from ISIS have stated that they target young individuals because they oftentimes do not have a concrete sense of self belonging (Awan, 2017). Those struggling with self-identity and the desire to belong to an organized group also showed strong levels of vulnerability to online extremism (Schils; Verhage, 2017). There have been several instances where young individuals have been caught by the FBI at airports or in the process traveling to join an extremist organization. When security officials have further investigated these individuals, they often find correspondent history between young adults or teenagers with terrorist organization that claim to offer belonging with their group (Engel et. all, 2016).

While Schils studied why people are attracted to oline extremism, Laura Huey looked specifically at how extremist messages are portrayed to young individuals on social media platforms such as Facebook, Ask.fm, and YouTube (Huey, 2015). Political jamming is a tactic used by terrorist that takes violent content and creates a narrativ e portraying extremist activity as "cool" or "humorous" in order to attract young audiences. Though it was argued tha t political jamming and online extremisms are not directly

related to radicalization, they do have a normalizing effect on political violence which is particularly dangerous to developmentally vulnerable adolescents (Huey, 2015). Huey used Twitter in her study by analyzing and documenting several tweets that had been politically jammed by pro-jihadists (Huey, 2015). She found that much of the pro-jihadist's material was aimed at adolescents in Western culture. Her study emphasized the importance of understanding how society receives information in order to better counter extremism efforts through methods that even the most vulnerable individual will connect with (Huey, 2015).

Costello's research contributed to Huey's evaluation of how online extremism is portrayed by looking at the specific behaviors and attitudes which led young individuals to extreme content (Costello, 2016). He asked if a lack of trust in the government led to increased exposure and, overall, what behaviors made youth and young adults more at risk for viewing online extremism. By collecting survey data from panels of individuals ages 15 to 36, it was found that over 65% of the participants were exposed to hate material, and most individuals were accidentally exposed (Costello, 2016). Costello also found that exposure to extremist content was heavily related to online behaviors such as search history and the amount of time spent online. He argued that, at this point in time, right-wing extremism is the most prevalent online; thus, individuals who searched content that was conservative in ideology or anti-government in tone had a greater likelihood of encountering extremist content (Costello, 2016). Both an unrepresentative participant age range, as well as ideological discrepancies posed limitations to this study. However, the findings were still valuable, as they help explain who is the most vulnerable to encountering online extremism (Costello, 2016).

Rather than analyzing the victims of online extremism, Philip Baught and Katharina Neumann researched how right-wing extremists use and perceive the media, in addition to how they attempt to bait media coverage (Baught; Neumann, 2019). By conducting interviews with former leaders of right-wing, extremist groups in Germany, they asked how leaders' perception of the media influenced their extremism, and how they described the emotional and behavioral effects of inaccurate coverage on news reports. The findings of the interviews and surveys showed that the media bears emotional and psychological effects on terrorists. In turn, this contributes to the overall danger of online extremism, since extremists become increasingly aggressive when they feel they are being inaccurately depicted by the media (Baught; Neumann, 2019). Similar to Huey's argument that counter-terrorism is dependent on understanding the culture of those most vulnerable to extremism, Baught and Neumann's research argued that it is equally as crucial to understand the goals and attitudes of extremists when the media are reporting. The final recommendation was to publish news in an unemotional manner by reporting raw facts only (Baight; Neumann, 2019). This study was limited, as only a small number of interviews were conducted with interviewees of Germany ethnicity only, which is not reflective of the behaviors of all extremist leaders. However, the findings were comprehensive, as desires for the media to portray German extremist groups in a threatening manner were most likely similar to how other terrorist organizations wanted their groups to be depicted (Baught; Neumann, 2019).

Elizabeth T. Harwood's research aligned with Baught and Neumann's emphasis on the importance of understanding extremists' behavior. She contributed to the research by comparing the manifesto of the 2019 New Zealand shooter to various political celebrities who use new media platforms to further their agendas (Harwood, 2019). It was found that the language of the shooter's manifesto was eerily similar to the language of political celebrities' speeches that were posted on social media sites, such as YouTube and Facebook (Harwood, 2019). This research suggested that more studies should be conducted on how political celebrities have the ability to influence the public in a radicalizing way even though they may not be radical themselves (Harwood, 2019). While Harwood advocated for celebrities to show greater responsibility when posting online, Meagan Flynn's (2019) article advocated for the public to also show responsibility by reporting extremist behavior on the Web. Flynn recorded that the livestream video of the New Zealand shooting was viewed over 4,000 times, yet no one reported it to Facebook. Harwood and Flynn's articles were not an outright call for censorship, but they did suggest that a policy should be implemented to minimize the opportunity for the public and potential radicals to view violent or strongly opinionated content that could inspire them to conduct an attack.

Censorship

One of the most contested policies in combating online extremism is the censorship of media platforms. Many scholars have agreed that censorship has flaws, and one of the first arguments against the use of censorship in the United States is the First Amendment's freedom of speech. Using professor of terrorism Paul Wilkenson's notes, Ubaysiri (2014) discussed how even though free speech has boundaries, censorship should be avoided because it erodes democracy and gives the enemy what it wants – uninformed citizens who doubt the legitimacy of the government. Additionally, Aaron Smith conducted research and found that many Americans already believed they were unfairly censored by media companies which resulted in a distrust for companies and for the government to intervene and protect citizens' free speech rights (Smith, 2018). In an international context, Article Nine of the United Nations' Universal Declaration of Human Rights states that the freedom of expression is recognized as a collective human right to all individuals (Wu, 2015). In addition to the freedom of speech, Anderson Terry (1993) argued in his article, "Terrorism and Censorship: The Media in Chains" that it is the right of the people to be informed by the media, and it is the duty of the media to provide

information on global activity to the public. Censorship is mostly rejected in the United States due to the protections that the Constitution provides in regard to free speech. Though scholars have argued that free speech is crucial to democracy, there is also a far greater risk for terrorists to abuse the accessibility of the Internet. The National Consortium for the Study of Terrorism and Responses to Terrorism provided data from 2005-2016 that showed a 78% increase in social media usage by terrorists in the United States (START, 2018).

Though many scholars have claimed that censorship is a violation of free speech rights, others have disagreed and suggested that the First Amendment and Declaration of Human Rights do not apply without boundaries. Melissa J. Morgans and Victoria Killion discussed the various forms of unprotected speech, thus arguing that the language which terrorists and extremists use on the Internet is comparable to the threatening speech that is not protected under the First Amendment, thus, justifying censorship (Morgans, 2017; Killion, 2019). In the United States, the government has had the ability to limit anti-government speech through censorship during times of war (Morgans, 2017). Morgans preceded her argument with the claim that the U.S. government should have the right to censor extremist speech online since there has been consistency in the antigovernment tone of terrorist propaganda. During times of war, the United States government has censored media content that contains disturbing or shocking images related to war (Morgans 2017). Morgans stated that this type of censorship also correlates with terrorists' use of the media since extremists often post graphic images or videos of violent activity that elicits a great sense of fear (Morgans 2017).

Along with support for anti-government and hate speech censorship, other scholars have found that the public shows an overwhelming support for censorship of sexually violent content (Fisher; Cook; Shirkey, 1994). Researchers Cook, Fisher, and Shirkey conducted a survey with both male and female individuals, finding that the majority of their participants supported a total ban of sexually violent movies, videos, and magazines (Fisher; Cook; Shirkey, 1994). Though censorship of sexually violent content differs from that of extremist content, one can deduce a broader desire of the public to prevent violent content, no matter what context it is in. One of the main limitations of censorship agreed by many scholars is its effectiveness. According to David Lowe's article, the first step to implementing counter-extremism measures, such as censorship, would be to find a universal definition of what constitutes "extremist" content (Lowe, 2017). Each country within the United Nations holds a differing definition of terrorism and extremism, thus creating difficulty in developing a parameter to monitor and censor extremist content. Another limitation for censorship would be the expansion of the Dark Web. According to Weimann's article, the Dark Web is approximately 400-500 times larger than the surface web and consists of illegal material that cannot be accessed through general Web browsers (Weimann, 2016). He claimed in his article that social media censorship of hate propaganda is a main contributor to the increased usage of the Dark Web, since terrorists and other extremists can post their content without repercussions (Weimann, 2016). Though, as Dow Jones Institutional News (2017) reported, the Dark Web had been cracked by the FBI in past years; however, more funding and technology would be necessary to consistently rely on national security to uncover Dark Web users.

Counter-narrative Programs

Other scholars have rejected the notion of censoring online content and, instead, argued that free speech should be used as an educating and mentoring tool to combat Internet extremism. Peter Neumann discussed several counter-narrative possibilities for the government to promote and evaluate programs that have already been implemented (Neumann, 2013). He argued that awareness of online extremism is a crucial start to prevention, suggesting that the government should be more vigilant to adopt digital learning platforms that adolescents and young adults can use to better understand how to navigate the web (Neumann, 2013).

Similar to Neumann's suggestion of digital education as the strongest preventative approach, Sarah McNicol discussed two methods that the United Kingdom implemented in schools to protect youth from online extremism – online filtering and digital literacy (MnNicol, 2016). She conducted a survey and found that only 9.4% of respondents in the United Kingdom disagreed that the Internet should remain unfiltered in schools, thus concluding that Internet filtering is overwhelmingly accepted in the U.K. Though censorship had been accepted by much of society in the U.K., many school librarians argued that research was severely limited for students, causing a deficiency in social abilities and development (McNicol, 2016). Interviews with school librarians provided evidence that Internet filtering actually caused a falsified sense of security, and students believed that the sites which remained unblocked were always legitimate. Five groups of students from the ages of 11-17 were interviewed in McNicol's study. The majority of students said that they were unable to gather the information necessary to complete assignments because the topics of the assignments were often filtered due their

controversial content, such as abortion and drug abuse. Additionally, students claimed to understand the dangers of the Internet, but said teachers did not emphasize the importance of learning how to navigate the Internet safely (McNicol, 2016). McNicol concluded the article by recommending digital learning programs to be required in schools so students could better understand the workings and risks of the Internet.

While McNicol and Neumann both argued that digital literacy and counter-narrative programs are effective policies in combating online extremism, Alex Wilner and Brandon Rigato discussed various counter-narrative programs and analyzed their limitations, along with their overall effectiveness. The study evaluated a sixty-day Preventing Violent Extremism (PVE) program that was conducted virtually in Ottawa, Canada. Wilner and Rigato claimed that most PVE campaigns were narrowly targeted towards Islamic extremism, often ignoring right-wing and other forms of extremism (Wilner; Rigato, 2017). The PVE in Canada focused equally on all forms of extremism in the program, developing a website that would produce counter-narrative content, information on where to find credible news sources, tools and resources for individuals struggling with radicalization, and graphics that provided information on current events (Wilner; Rigato, 2017). The site gained a lot of traction with 1,500 likes, over 550 shares, and 1,700 views. The program seemed to be a success in the attention it garnered. However, Wilner and Rigato argued that the quantitative data were not accurate reflections of efficacy, as it could not be measured whether the posts contributed to an increase in viewers' knowledge of extremism, or if the program was responsible for preventing an individual from radicalizing (Wilner; Rigato, 2017). Additionally, Logan Macnair and Richard Frank discussed a case study on a counter extremism campaign called Voices Against Extremism. In the United Kingdom, Prevent has been a counter-terrorism strategy which seeks to avert radicalization and the promotion of online extremism (Macnair; Frank, 2016). Macnair and Frank discussed several flaws with Prevent and various Counter Violent Extremism (CVE) campaigns. Similar to Wilner's findings, they argued that CVE campaigns lacked the ability to be measured in effectiveness since radicalization and extremism do not have a universal definition or scale of intensity (Macnair; Frank, 2016). Additionally, many CVE initiatives focused on creating national policies, and the authors argued that a community-based or individualistic approach would be a better, more effective alternative (Macnair; Frank, 2016). Since Wilner and Mcnair argued that online count-narrative programs are limited in their ability to be empirically measured, in-person deradicalization could draw a parallel to the effectiveness of online programs. Samantha Kutner discussed a deradicalization program in Saudi Arabia which has worked with extremists in different categories (Kutner, 2016). Each committee focused on getting participants to willfully reject violent extremism and ideology, in addition to teaching participants how to help other extremists deradicalize. Though the Saudi government has claimed to have high success rates in its program, it is not completely evident that the program has such a high success rate, as more longitudinal research would need to be conducted on extremists' post-rehabilitation (Kutner, 2016).

Media Accountability

The final policy analyzed in this thesis is media accountability for extremist content that Internet platforms host and allow to be produced. Softness (2016) and Hayden (2019a) argued that social media companies such as Twitter, Facebook, and Google have failed to uphold their pledge to better monitor and control their platforms against the spreading of extremism. While some companies have ignored warnings from various platforms on potential extremism existing on their site and faced little to no repercussions, others have been sued for responsibility for terrorist attacks. In 2016, Google, Twitter, and Facebook were sued for responsibility for the 2015 Paris terrorist attacks, since the companies were allegedly knowledgeable that their platforms were being utilized by terrorists. Softness claimed that the government has failed to develop effective legislation which governs online activity in a timely manner and continued by arguing that holding Internet owners accountable for extremist content would create online stability (Softness, 2016 & Romm, T., Harwell, D. 2019). A current limitation in holding companies accountable can be found in the U.S. 1996 Communication Decency Act, which says that jurisdiction can only be held against the extreme content, not against the platform the content is hosted on. Additionally, it was argued that media accountability requires proof of correlation between the extremist content and physical actions of extremism. Softness countered these limitations with U.S. Code 2339(A) that puts liability on persons that provide or facilitate material support for terrorists (Softness, 2016). She argued that Internet platforms who host extremist content are providing material support to terrorists since their platforms act as a tool to spread propaganda and recruiting content.

Alexander Tsesis discussed the limitations accompanied with holding media companies accountable for online extremism. The Communications Decency Act in the United States has granted immunity to service providers, search engines, and social networking sites, but the First Amendment has allowed Congress to impose liability on service providers who are aware that their platforms host terrorist activity (Tsesis, 2017). Tsesis argued that an enforcement of material-support laws would be the most effective way to combat terrorism online since companies would be held liable for allowing extremist content to remain on their platforms. Similar to the limitations discussed in Softness' article, there would need to be proof of causation between extremist content on Internet platforms and a physical action of violence in order for the case to be determined as valid in the courts. Under strict scrutiny in the court, the plaintiff would also need to prove that she was legitimately harmed by the content posted on an

Internet platform. In addition, evidence would need to prove that the media company at question had knowledge of the extremist content on its platform but failed to regulate it (Tsesis, 2017). Even if there were evidence that the company was aware of the existing radical content, the Counter-Extremism Project (2017) counter-argued that it is within the rights of media companies to monitor their platforms according to their Terms and Conditions policies – even if this results in extremism remaining online. Not only does media accountability have legal difficulties, there are ethical problems in implementing strict repercussions on the media to better monitor its platforms. In Terry Gross' (2019) NPR discussion on content moderators, the moral controversies of the position were brought to light since individuals who have worked as moderators faced extremely graphic and disturbing content that could contribute to psychological trauma or radicalization.

Emily Tate contributed to the legal aspect and difficulties of enforcing media accountability in her article. She argued that, since individuals who provide material for domestic terrorists receive no civil remedy, the federal government has limited capabilities in holding regulatory authority over companies of media platforms (Tate, 2019). She called for an amendment that would include material support to both domestic and foreign terrorists, arguing it would be narrow enough to be effectively applied, while ensuring lawsuits still had high standards requiring proof that edge providers gave material assistance to extremists. Tate argued that this new implementation would encourage better practices by Internet providers. To add to the legal analysis of media accountability, Jordy Krasenberg (2019) discussed the European Union's strict efforts to prevent online extremism in its member states. After the New Zealand shooting's live stream footage was spread internationally on social media outlets, the EU implemented a strategy that would more heavily regulate terrorist content online. This would require member states' media to remove terrorist content within one hour of its original post or face strong financial penalties for noncompliance (Krasenberg, 2019). Additionally, proactive measures would be required of member states. Each state's media companies must invest in technology that assists in detecting and removing terrorist content and increasing the blocking of illegal content (Krasenberg, 2019). Companies are also subject to annual transparency reports by the EU to increase accountability and ensure compliance with the new Internet standards. Russia has enforced similar regulations on media outlets. Australia has also tightened its measures with fines or jail time as consequences for companies who fail to quickly and efficiently moderate their platforms (Conifer, 2019). In Sharyl Cross' journal, Russian extremism was discussed at length, and addressed how the Russian Security Council created a "watchdog" system that would provide formal warnings to Internet providers to remove extremist content from their platforms (Cross, 2013). The policy implementations were still new upon the creation of the article; thus, more longitudinal studies would need to be conducted to analyze its efficiency. Overall, the countries mentioned depict an urgency and seriousness that international governments have taken in recent years to address online extremism.

With the increasing capabilities and accessibility to new media outlets, the discussion of online extremism is becoming more crucial to the safety of everyday Internet users. The existing research on this topic is beneficial in understanding what online extremism is, who it affects, and what policy options exist in combating new media extremism. However, each policy option must be empirically measured to test whether it is both pragmatically and normatively feasible in dissuading online extremism.

III. Methodology

This thesis will examine the three categories of policy options that are most commonly debated among scholars: censorship, counter-narrative programs, and media accountability. Since online extremism is a recent problem for various governments and there are limited data on the efficacy of the suggested policies, this thesis will primarily focus on the normative and pragmatic complexities involved with each. By analyzing a mixed variety of data through case studies, anecdotal evidence, legal codes, and historical precedent, the thesis will conclude with a summary and general recommendation on which option would be the most normatively and pragmatically feasible in mitigating online extremism.

To analyze the normative success of censorship, this thesis will examine various cases of historical precedent in the United States that have been used during times of war. Additionally, the First Amendment of the Constitution will be used to address the legal complications of enforcing extensive measures of censorship within the United States. After discussing the normative aspects of censorship in America, this thesis will explore how international governments have created censorship laws in response to online

extremism, then address the effectiveness and limitations of these legislations in Russia and the European Union. To address the pragmatic aspects of censorship, data on the public opinion of censorship for sexually violent and child pornographic content will be examined, then used to draw a parallel to the challenges associated with censorship laws to address online extremism. This paper will also discuss the use of the Dark Web by extremists and terrorists who are censored, then deliberate the limitations of censorship due to the use of the Dark Web.

One of the main limitations of online counter-narrative programs is their difficulty to be empirically measured. Several underlying factors can contribute to an individual's deradicalization process, thus making it nearly impossible to give a specific statistic as to what level the counter-narrative program actually contributed to the decrease of extremism. However, face-to-face deradicalization programs have similar goals as online counter-narratives. To give a better idea of how such programs operate both normatively and pragmatically, this thesis will analyze Saudi Arabia's deradicalization program, then draw a parallel to address how effective online counter-narrative programs could work. In addition, national security threats and ethical dilemmas associated with specifically targeting Islamic terrorist organizations will be discussed briefly.

To address the normative logistics of media accountability, this paper will analyze the strict laws that the European Union has placed on its member states' media outlets and compare them to the 2019 EU terrorism trend report. The U.S. Code § 230 of the 1996 Communications Decency Act – which protects media platforms from accountability for content they host – will also be used to address the complexities of the legal implications for media accountability in the United States (Softness, 2016). To measure the pragmatic potential for media accountability, I will assess the United Nation's dilemma in addressing online extremism, then compare it to the EU's Code of Conduct and E-Commerce Directive; this requires media platforms to comply to a strict set of regulations in monitoring and removing extremist content (Wu & Softness, 2015; Europol, 2019).

IV. Definitions

For the scope of this thesis, the definition of online extremism will be based off of the Counter Extremism Strategy of 2015. The project defined extremism as, "the vocal or active opposition to our fundamental values, including democracy, the rule of law, individual liberty, and respect and tolerance for different faiths and beliefs. We also regard calls for the death of members of our armed forces as extremist" (Educate. Against. Hate., 2020). To make the definition more tailored for this thesis, hateful or intolerant speech towards race, gender, and religion, in addition to posting disturbing or violent content in order to evoke fear, hate, or radicalization will also be included in the definition. The purpose of this thesis is to explore both the normative and pragmatic limitations and successes of each policy. Normative will be defined as any legal, constitutional, or ethical prospects of a policy. The definition of the pragmatic prospects for this thesis will entail the practical and realistic aspects of each policy. Censorship will be defined as the blockage or removal of any content that is included under the extremism definition or is a threat to national security. The Dark Web is defined as sub-surface levels of the Internet that cannot be accessed by normal search engines. The Dark Web allows users to remain anonymous, making it easier for them to conduct illegal activity such as child exploitation, money laundering, the facilitation of communication between terrorists, and allows extremist content to be posted without censorship. Counter-narrative programs are defined as online programs which seek to offer a counterpoint to extremist content and attempt to have a deradicalizing effect on an individual. To draw a parallel between online and in-person counter narrative programs, Saudi Arabia's deradicalization program will be used as a model. According to the Journal for Deradicalization, the concept of deradicalization is defined as, "the social and psychological process whereby an individual's commitment to, and involvement in, violent radicalization is reduced to the extent that they are no longer at risk of involvement and engagement in violent activity" (Kutner, 2016). Media accountability is a straight-forward definition entailing any legal repercussions and responsibilities that are placed on new media platforms to hold them accountable for the content they host.

V. Results

Policy I: Media Censorship

The first policy examined in this thesis is media censorship. As mentioned in the Literature Review, in the United States, the First Amendment's rights to free speech and a free press has made it difficult for legislators to constitutionally implement censorship laws. However, while the Constitution protects speech, this freedom is not completely unrestricted. According to the Federation of American Scientists, threatening speech, obscenity, fraud, child pornography, speech that encourages illegal activity, fighting words, and defamation with the intent to cause injury are types of speech that the Constitution does not protect (Killion, 2019). Normatively, this could give legislators some flexibility in implementing censorship legislation since much of extremist content consists of these categories of speech.

Throughout the history of American politics, censorship and limitations on the freedom of speech have been temporarily implemented during times of war. Speech that is anti-government in language or that opposes basic American ideals has been blocked by the government at various times (Morgans, 2017). Additionally, images that showed the violence of war or depicted America as the weaker and losing side of a war have also been banned during times of conflict. Some examples of historical censorship can be seen in the Revolutionary War when various Loyalist newspapers were stolen by the Patriots in order to prevent information from being dispersed to the public. Additionally, during the Civil War, President Lincoln

ordered telegraphs to be seized in order to disrupt Confederate communications. Artists who painted or drew gruesome photos of the war were told to "tone down" their imagery to avoid causing fear or doubt towards America's military capabilities amongst the public (Morgans, 2017). Graphic photos were also censored during World War I, and the Espionage Act of 1917 made it a crime for individuals to promote or display propaganda of the success of American enemies. Even throughout the Cold War, there were restrictions placed on the press and media platforms, including the criminalization of promoting anti-government rhetoric. More recently, after the September 11 terrorist attacks, former President George W. Bush signed the USA Patriot Act, which gave greater abilities for the government to monitor and restrict both the press and the public's free speech during times of conflict (Morgans, 2017). Normatively, it can be argued that, according to historical precedent, extremist content should be censored since it is predominantly anti-government in language and promotes gruesome propaganda that causes a disturbance to the public. There have been numerous instances of ISIS and Al-Qaida posting execution videos and intimidating images of foreign fighters holding hostages at gunpoint with the purpose of evoking fear in a mass audience. They also have been known to use rhetoric which appeals to various audiences through anti-government language that resonates with those who have shown animosity towards the government. Oftentimes, ISIS and Al-Qaida have used narratives that depicted their perspective of the evil of America in attempting to interfere in Middle Eastern affairs and called for the overthrow of America as the leading world power.

Many White supremacists have also used unprotected forms of speech in their manifestos that openly threatened and called for violence against minorities such as Muslims, Hispanics, and African Americans. In the recent El Paso shooting, the shooter had posted a manifesto to social media platforms that was filled with anti-immigrant language and a call for violence against Hispanics, referring to immigration as a "Hispanic invasion of Texas" and a "Great Replacement", which is the belief that the White population in the West is being overtaken by minorities (Hayden, 2019b). Calls for action to forcibly stop immigration through violence are blatant threats toward a specific population, which give such speech Constitutional grounds to be censored since threatening speech is not protected under the First Amendment. There could be a strong argument that the censorship for such content is not only feasible and legitimate but is also necessary to mitigate online extremism in the United States.

Though there has been historical precedent and legal abilities for the government to use censorship as a tactic in mitigating extremism, there are also normative complexities that must be addressed before enacting censorship as a core policy. The Counterextremism Project argued that censorship of extremist content is a Terms and Conditions issue rather than a free speech debate, since companies such as Twitter, Facebook, YouTube, and WhatsApp are private entities capable of censoring content that categorizes under the "extremist" label (Counterextremism Project, 2017). For many years, several social media companies have consistently shown hesitation in censoring extreme content, fearing that those who are blocked or removed would sue for "free speech" violations which, in turn, would damage the company's reputation and overall profit. Laws in the United States have also prohibited individuals from granting material support and technological advice to terrorists through the media, yet platforms have continued to abstain from removing users or content that poses a clear threat to security (Counterextremism Project, 2020). Many social media platforms have ignored warning signs received on their sites from Hatewatch, a platform dedicated to monitoring and flagging extremist material (Hayden, 2019a). The Southern Poverty Law Center claimed that Hatewatch had flagged Google three times to inform them of extremist content hosted on Telegram (an app Google provides to all of its consumers); however, the company failed to respond and continued to provide the application despite the knowledge of existing extremism (Hayden, 2019a). Accordingly, legislators have begun tightening the criteria for content that is recognized as extremist or helpful to terrorists, which has brought some large social media networks under scrutiny in recent years. In the USA Patriot Act sections 805(a)(2)(b), the definition of material support was given a broader meaning due to the recent criminalization of providing terrorists with any kind of support or advice – not just monetary or technical assistance as before. Additionally, since 2010, in the Supreme Court case Holder v. Humanitarian Law Project, hate speech could also be considered a crime if it showed support of, or coordination with a foreign terrorist group (Counterextremism Project, 2017). In October of 2018, representatives from Facebook, Twitter, and YouTube were brought before Congress to testify on their efforts to prevent users from being exposed to extremist content online

(Romm & Harwell, 2019). Each platform claimed to be successful in improving the overall security of their sites, even claiming that 90% of nine million extremist videos, channels, and comments had been removed by automated technology on YouTube (Romm & Harwell, 2019). However, Congress argued that their efforts were insufficient due to the massive volume of extreme content that had still been posted on each platform, then viewed and reposted. For media companies to better censor their platforms, an argument can be made that Congress must enforce tighter rules on the Internet and various online companies.

From an international perspective, the Russian government has aggressively combatted online extremism since the Arab Spring protests and expansion of social media within the country from 2011 onward (Cross, 2014). The Russian Federal Agency (RFA) bears the responsibility of monitoring the Internet and news media for "objectionable" content. Once such content is detected, the agency has a twenty-four-hour period to remove the content from the Internet or face fines and suspension of service (Cross, 2014). Additionally, in 2012, Russia's Mass Media Inspection Service was granted the ability to censor questionable content without the need for a court to rule in favor of the blockage of detected content. Russia also has a Federal Law that combats online extremist activity but has received strong amounts of criticism for its ethical implications. According to Cross, the law "does not require establishing the threat of the use of inciting violence for prosecution (Cross, 2014, p. 13). This means that even peaceful groups (typically religious) can be targeted as "extremist" for their beliefs that do not align with the majority or the norm. Additionally, between 2007 and 2013, over 1,500 writings were banned from the Internet and labeled as "extremist" – most of which were Islamic literature and religious writings (Cross, 2014). Though the laws that Russia has implemented have been accepted by much of the population, there are ethical and legal debates that can be used against the tactics that the Russian government has utilized in combating online extremism. Especially in the United States, where there are constitutional protections to minorities and all creeds, such strict and discriminating legislation would undeniably fail to pass through Congress. Additionally, there is limited data on whether the measures that Russia has taken in censorship have actually helped to reduce the amount of extremism on the Web. Cross stated that one of the main normative challenges of Russia's Internet security is the need for a global response to address online extremism. Another legal challenge that Russia's situation presented is the need for all countries to develop protocol on what measures of Internet censorship would be ethically acceptable and legally feasible (Cross, 2014). The vast differences in national values, laws, and historical precedent make it extremely difficult for various countries to compromise on developing a universal standard to mitigate extremism.

The European Union has also tried to enforce stricter laws upon its member states to curb extremist or terrorist content from reaching their social media platforms. In the EU, extremist content and hate speech is illegal, but since social media faces no international borders, member states have still encountered instances where content was posted from another country and spread online to Europe, which has caused the EU to tighten its laws even further. Prior to the 2018 New Zealand mosque shooting, Article 15 of the EU E-Commerce Directive protected service providers from having to monitor and censor their platforms against extremist content (Softness, 2015). Now, the EU member states are legally required to efficiently censor their platforms and block extremist content from being posted or face financial repercussions. The EU E-Commerce Directive was amended in 2018, making it too early to tell if the new regulations have a significant effect in censoring social media platforms. However, according to the Terrorist Trend Report of 2019, the EU made some successes in hindering ISIS's online presence despite their propaganda videos, advice on how to avoid detection and deletion, and an aggressive call to support ISIS media (Europol, 2019). Legally, censorship seemed to have some successes in the EU with the decline in the presence of online terrorist organizations; however, it is important to note that it is difficult to measure if the decreases of incidents are strictly due to censorship efforts, as several other factors could have contributed to the decline.

Though there are normative opportunities for censorship to be implemented, practically, there could be strong arguments to suggest that censorship would be a complex policy option to enforce. The first pragmatic challenge of censorship is the fact that the Internet will never be fully secure. Even if social media companies banned a user from their platform, that user could still have the capability to create an alternate account under a different name and continue to spread propaganda. In "The Efficacy of Censorship as a Response to Terrorism", Hezbollah's television network, Al Manar was banned by the US State Department and coined as a terrorist organization. Even with the censorship of Al Manar, the network was still available to areas of Europe, North Africa, and the

Middle East (Ubayasiri, 2014). Due to the increasing variety and availability of social media platforms, some scholars have argued that censorship efforts will hardly decrease online extremism.

The use of the Dark Web would be another practical limitation of censorship, as users who have been banned would have the option to post their extremist content on an anonymous platform. As discussed in the definitions section of this thesis, the Dark Web refers to any Internet content that is not reachable by general search engines (Weimann, 2016). Though there is no firm data on the actual size of the Dark Web, some researchers have argued that it is roughly four-to-five hundred times larger than the regular Internet. Among pedophile activity,
the buying and selling of illegal drugs and weapons, and credit card fraud, terrorists also have been able to maintain a strong presence on the Dark Web (Weimann, 2016). The Dark Web has been a beneficial resource for terrorists since they are able to fundraise, recruit, communicate, and gather data with little to no fear of being uncovered by the government.

Practically speaking, censorship of extremist or terrorist content on the Internet would most likely lead to an increased presence of terrorist activity on the Dark Web, which could be more dangerous than a presence on the general Internet. Since Dark Web users are anonymous, it is significantly more difficult to track terrorists and extremists. In Cross's (2013) article on the Russian responses to terrorism, she made the argument that extensive censorship efforts could be considered as counter-productive for cyber security since users are not anonymous on the general Internet, and it is easier to monitor and surveil potential threats.

However, there have been some instances where the Dark Web has been cracked by the FBI and a noted University, Carnegie Mellon. In 2017, two marketplaces that sold criminal goods such as firearms, heroin, and fentanyl were shut down and the marketplace owners were arrested (Dow Jones Institutional News, 2017). The site had serviced around 200,000 users with approximately 1,000 sales per day. Attorney General Jeff Session argued that cracking the Dark Web had been the most important task of the year for the FBI, and that it would cause Dark Web users to be wary of using the site. If the government was able to invest more finances into technology to crack the Dark Web on a more regular basis, it could be argued that censorship would be more of an effective method in eliminating extremist content since (due to the higher risk of being caught) terrorists and extremists would be less likely to use the Dark Web. The Weimann article argued that, in order to get the technology necessary to combat the Dark Web, it would be crucial to provide consistent and pressing evidence of the seriousness of the Dark Web as a primary platform for crime and global terror (Weimann, 2016). The Defense Advanced Research Projects Agency (DARPA) conducted a long-term study on the workings of the Dark Web, and suggested a software called MEMEX (Memory and Index) that would provide more efficient cataloging of the Dark Web and thereby penetrate the site and uncover those involved in anonymous activity (Weimann, 2016). Pragmatically, for censorship to be an effective policy in removing extremist content from both the regular Internet and the Dark Web, software with the capability of consistently cracking and uncovering the Dark Web would be an essential tool for policy makers.

As mentioned in the Literature Review, the public opinion on censorship in the United States has generated a mixed response based on the content at question. Sexually violent content and child pornography is argued to be deserving of censorship due to its graphic, taboo, and threatening nature. However, Americans' opinions have been inconsistent in the view of what constitutes content that warrants censorship, even though the majority of extremist content is of similar nature, albeit with different origins. In areas of free speech, the public has been more hesitant and skeptical of censorship. According to PEW Research, 72% of Americans believed that social media companies actively censored any political content that did not align with that platform's political leanings and argued that any censorship is an infringement on basic American rights (Smith, 2018). Additionally, PEW recorded that only 3% of Americans trusted that media platforms would "do the right thing" when confronted with a situation that required action, and roughly half of the public believed that these media companies needed more regulation to ensure they were not actively censoring any forms of content regarding speech (Smith, 2018). From a practical standpoint, the American public opinion of censorship matters greatly when legislatures or private companies enact policies that would remove extremist content. Since legislatures rely on the American people for re-election and companies rely on consumers for revenue, the ability to enforce stricter censorship policies is risky. Additionally, the perceived legitimacy and credibility of the government oftentimes depends on the public approval of the various laws and regulations that are implemented. In order to maintain

credibility and trust among the American population and further the emphasis for democratic values, the government must be cautious and intentional when considering sensitive censorship policies.

Policy II: Counter-narrative Programs

The next policy under consideration is online counter-narrative programs. As aforementioned, online counter-narratives have been difficult to empirically measure due to the uniqueness of each individual's degree of radicalism. Normatively, one area of concern is the targeting of the Muslim population in such programs. Though recent counter-narratives have become more diverse in the forms of terrorism they address, there has still been a far greater emphasis on Islamic extremism over left-wing, right-wing, or any other forms of terror. As Wilner and Rigato (2017) discussed in the summary of their 60-day counter-narrative program, it is crucial for such programs to equally address all forms of terror. Not only does targeting Islamic extremism lead to higher levels of discrimination towards the Muslim population, but it also plays into the fear-evoking narrative that legitimate terrorists want the populous to be paranoid of. In the United States, the First Amendment grants the Constitutional freedom of religion to all creeds; however, since the September 11th terrorist attacks, a sharp influx of prejudice and hate towards Islam has occurred, resulting in the generalization of Muslims in counter-narrative programs.

Targeting Muslims in counter-narrative programs could also contribute to national security threats. For the United States to effectively combat jihadi terrorists and extremists both on the Web and in real-world scenarios, human intelligence from the Muslim community is crucial. By discriminating, alienating, and targeting Muslims with these programs, it is likely that moderate Muslims would become less likely to cooperate with national security personnel. Editor of Politico Magazine, Michael Hirsh found that a Muslim community in Dearborn, Michigan was willfully engaged with national security agencies such as the FBI and Department of Homeland Security due to the outreaches and resources that the community provided for them. The resulting mutual trust and respect between the authorities and community led many of its residents to express their willingness to – and, for some, their fulfillment of – reporting radicalized family members, peers, and friends (Hirsh, 2016). In contrast, Hirsh interviewed Muslims from Europe who said they would be less likely to cooperate with the government due to feeling targeted and discriminated against. Arguably, the best method for encouraging individuals to assist national security agencies is to promote equality and trust between the public and authority. By building trust and providing assistance to the public, mutual respect and commonality in community goals would increase. Targeting and surveilling a certain group based on their race or religion is counterproductive and dangerous to counterterrorism efforts. For counter-narrative programs to be successful, they must be based on factual evidence of prevalent threats and avoid any stereotypes or prejudice towards specific groups.

Recent data has shown that domestic and right-wing terrorism is currently a larger concern in the United States than Islamic extremism. According to 2018 data collected by the FBI, rightwing terrorists outnumbered Islamic terrorists by a ratio of 120-to-100 respectively (Stieb, 2019). Right-wing terror has also posed a significant threat throughout Europe. The Policies for decrease of terrorism mentioned previously by Europol was mostly due to a decrease in ethno-nationalist and right-wing terror since this was the most predominant form of extremism at that time. Europe experienced a sharp influx in right-wing terror in 2015, when each member state of the European Union was required to follow migrant quotas and increase the number of refugees allowed in the country. Both strategically and practically, counter-narrative programs should shift their focus to the largest threat that current trends are illuminating.

Pragmatically, considering counter-narrative programs as a core policy in mitigating online extremism could be argued as problematic. Though there has been an increase in the amount of time society and extremists spend online, there is no guarantee that radicals and potential terrorists would visit these sites or further engage in the programs if they did encounter them. Even if they did, the chances of extremists permanently deradicalizing cannot be empirically measured, which is a crucial aspect in determining policy efficacy. The same concern of inability to be empirically measured can be seen in Saudi Arabia's Risk Reduction Initiative (RRI) program. The unique deradicalization tactics that Saudi Arabia has used involved four "committees" that each extremist must seek for help – religious, psychological, security, and media (Kutner, 2016). The goal of the program is to get participants to reject all violent ideology and world views that they have adopted. In addition, participants are offered financial and occupational assistance which attempts to help transition former extremists back into self-sustaining members of society. After participants finish the program, they are monitored

with surveillance to test the amount of recidivism after the program (Kutner, 2016). However, the surveillance that participants temporarily receive is difficult to measure recidivism rates; it is impossible to analyze a participant long-term, and he or she could return to radicalization upon the completion of parole. The Saudi government believes their programs to have an 80-90% success rate; nevertheless, the pragmatic critique of the program is that no component of the program has been studied for efficacy (Kutner, 2016).

Another practical difficulty of the Saudi program is its fundamental goal of reeducation. By the time an extremist becomes eligible for participating in such a program, their ideology, religious views, and thinking system have most likely already been engrained, making it extremely difficult to completely reeducate and alter their way of thinking. There is also an ethical concern associated with the reeducation process, in that the process by which psychologists work with participants must be a way that humanely retrains extremists' cognition. The question of efficacy and ethical conduct is even more prevalent for online programs as opposed to in-person reeducation. The ability to have face-to-face conversation with an individual allows greater and more impactful opportunities to connect with the extremist. Online counter-narrative programs would be difficult to implement due to the fact that each individual requires unique methods of reeducation, experiences varying degrees of extremism, and has diverse backgrounds that could alter the needs of the person. The fact that these programs have been so under-researched and do not have a concrete method of measuring efficacy arguably make it a weak policy to implement.

Though there are areas of growth and research that must be addressed in counter narrative programs, there are both normative and pragmatic areas of success that may warrant further interest in such policies. From a normative standpoint, there are little to no legal, ethical, or Constitutional problems with such programs (apart from overcoming the general discrimination against Muslim communities). Each counter-narrative follows guidelines that have been implemented by professionals who study extremists' behavior, then creates inclusive and informative platforms with the goal of reaching even the most radical of Internet users. Ethically, there are few areas of concern regarding counter-narratives. If the platforms promote factual evidence and create an environment that reaches all forms of extremism in a non threatening and humane way, these online alternatives to extremist content could show some successes. Practically speaking, programs are difficult to measure and impossible to ensure that every online extremist is exposed to – and participates in – such programs; however, if the overall goal is to utilize counter-narratives as a supplemental resource in decreasing radicalism, then they could be considered successful in themselves. Counter-narratives have the most potential for success if they are used in addition to other policies that focus on Internet security. Though statistical proof of success rates would not be obtained, placing counter-narratives online would still allow extremists the chance of encountering such a program and, perhaps, deradicalizing as a result. Nevertheless, using counter-narratives as a core policy in addressing online extremism would most likely be ineffective in dramatically reducing the amount of radicalization online.

Policy III: Media Accountability

The final policy that this thesis evaluates is media accountability. Normatively, media accountability bears notable limitations. Historically, the First Amendment in the United States has provided protection to corporations (in addition to citizens), making it difficult to enforce legal repercussions on media platforms for the extremist content they host. In addition to First Amendment protections, various codes in the United States have provided extra protection for social media outlets from absorbing liability for extremist content. Specifically, section 230 of the Communications Decency Act of 1996 states, "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider" (U.S.C. § 230). In past years, various cases have shown that courts typically rule in favor of corporations. A relatively recent example can be seen in the San Bernardino school shooting of 2015. The FBI had requested that Apple unlock the shooter's iPhone for investigative purposes; however, due to Apple's customer protection codes, they refused to unlock the device. The FBI sued Apple for refusal to cooperate and provide the information, but the court ruled in favor of Apple, arguing that the company was right in upholding its policies (Eidam, 2016). This case can be used as evidence of existing precedent in the United States to protect companies against legal coercion in revealing sensitive user information. Enforcing media accountability in the United States would require a major paradigm shift in laws to place responsibility on the provider rather than the producer of extremist content.

The freedom of the press is another constitutional covering that the government would need to navigate to effectively implement media accountability. A free press is crucial for any democracy to hold the government accountable for its actions and provide transparency for the public. But, as many scholars have argued, a free press also gives terrorists and extremists the publicity they need to generate fear amongst a mass audience. Ethically and constitutionally, holding the media accountable for reporting on terrorist and extremist activity could be problematic; it is historical precedent and a fundamental right of the American public to have access to national and international news. If legislation were posed against the press for reporting on extremist or terrorist activity, the public could fairly question the legitimacy of both the press and the government since these laws would radically shift the standard of a free press and limit the public's access to the news. However, a way to address this would be for the government to allow the press to report extremist and terrorist content but require reporters to use "self-censorship" tactics, and refrain from posting anything that would give terrorists extra publicity. (Anderson, 1993). Ethically and constitutionally, self-censorship could be a feasible option for the government to use as a tactic in addressing the problems associated with the free press; but, from a legal standpoint, self-censorship requirements would be nearly impossible to enforce due to the sheer volume of news reporters and an inability to regulate each outlet, article, or post.

From an international perspective, the implementation of media accountability is more complex and carries even further limitations. First, as Paulina Wu (2015) argued, implementing any kind of international legislation would require the United Nations to develop a universal definition of what the term "extremism" constitutes in order to enforce legal standards that extend to all member states. Within the United Nations, each country holds a definition of terrorism and extremism that is unique to their national interests. Some countries carry narrowly tailored criteria, and others use broad language that allows for subjective interpretation. For example, the country of France defines terrorism as actions that are, "intentionally committed by an individual entity or by a collection entity in order to seriously disturb law and order by intimidation or by terror" (Wu, 2015, p. 306). Also, the country's criminal code includes specific acts such as physical assault, kidnapping, the creating and selling of explosives or weaponry, and attempted murder as acts that are to be considered as terrorist behaviors. Other countries carry vague definitions of terror that make it difficult to enforce guidelines for accountability. The United States defines terrorism as, "premeditated, politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents" (Wu, 2015, p. 304). United States law also has specific characteristics of terrorism that include acts that violate criminal laws, intimidate or threaten civilians, and are intended to coerce or coerce the government. One of the critiques of the United States definition is its inconsistency (Wu, 2015). For instance, many shootings in the United States have been labeled as a hate crime rather than a "terrorist" incident, while other attacks performed with the same political, racial, or religious intent have been labeled as terrorist. When pertaining to the Internet, legislators would need to focus on the specific definition of extremism rather than terrorism; however, the diversity in definitions between France and the United States helps to illustrate the difficulty in reaching specific criteria for legislation. While some nations, such as the United States have valued a free press and the protection of citizens' speech, other countries would rather see tighter regulations on the media within their governments. With differing security goals and national values, it would be unreasonable to accrue a majority agreement on a definition of extremism and the appropriate legislation to enforce upon media platforms for failure to regulate their sites.

Not only does media accountability hold several legal, constitutional, and ethical implications, but it also bears pragmatic issues that warrant concern for overall effectiveness. Even if the United Nations was successful in developing universal parameters and legislation for media accountability and companies were effective in removing extremist and terrorist content from their platforms, it is most likely that censored extremists would resort to the Dark Web as discussed in the censorship section of this thesis. Additionally, the massive amounts of extremist activity online would be impossible for media companies to fully regulate. The Southern Poverty Law Center recorded that the social media app Telegram had gained approximately three million new users each day, making it impossible to effectively regulate each extreme post on that site alone – notwithstanding the thousands of other social media sites available (Hayden, 2019).

Practically, if media accountability legislation were enforced, there would not be enough content moderators to handle such large amounts of

censorship. Content moderators are hired by various social media outlets to screen and potentially block videos, photos, and posts that have been reported or flagged as inappropriate. According to NPR, content moderators oftentimes face severe Post Traumatic Stress Disorder due to the horrendous content they encounter up to thousands of times a week (Gross, 2019). On a day-to-day basis, moderators witness heavy amounts of content with pornography, child exploitation, terrorist beheadings, posts filled with hateful rhetoric, and blatant threats to various races and religions. Ethically, hiring individuals to work in such a psychologically disturbing position is controversial. It was also mentioned that, in some extreme cases, moderators who encountered extremist content at high frequencies began to believe the conspiracy theories and radical nature of the content they witnessed. An argument can be made that subjecting individuals to this type of work environment could potentially lead employees into radicalization due to the psychological trauma they face every day in their job. According to the research database RAND, the Internet behaves as a psychological reinforcer due to the echo chambers that users are exposed to. Since the Internet is an algorithm based off of the user's search history, a frequent Internet user is exposed to vast amounts of information that is specific to his or her online behaviors. With little diversity in content that shows alternate ideologies, one's current beliefs have been shown to be further ingrained into their cognition (Behr et. al, 2013). Similarly, individuals who are consistently exposed to extremist content are also at risk of having those violent and radical viewpoints infiltrated into their psychological behaviors and beliefs; this, in turn, may contribute to further radicalization, becoming counterproductive to counter-extremism efforts.

A content moderators' position is alarming to some who utilize the Internet since, as NPR discussed, many of the workers have not studied the Constitution or American Law, do not have college degrees, nor have lived in the United States for a long period of time. Many content moderators are brought to the U.S. from various countries and are unfamiliar with American politics, making it difficult to identify content that is censored by U.S. law. For example, while hate speech is legal and protected under U.S. law, it is illegal in many other countries, thus making it difficult for moderators to screen, as there are consistencies between hate speech and unprotected, threatening speech. If media accountability legislation were implemented, the power given to platforms would be similar to that of the government, as their responsibilities would entail the interpretation of the Constitution. Many could argue that it is neither the right nor duty of media companies to bear such legal authority in limiting speech.

Though there are many normative and pragmatic limitations to media accountability, there is also the potential for success in this policy option. In the United States, amending section 230 of the Communications Decency Act to include content that pertains specifically to terrorism could be an effective way to hold the media accountable while still protecting free speech. As argued in Wu's (2015) article, developing a tailored definition of extremism and terrorism would be crucial to accountability efforts. For example, the Communications Decency Act could be revised to include posts by a terrorist organization and content which contains blatant threats to a specific group, individual, or government organization. These alterations and inclusions could give media outlets greater guidance and accountability on how to address specific types of extremist content, while continuing to preserve other protections for media companies granted under U.S. law.

The EU's regulations on member states are also an example of how media accountability could produce positive outcomes. The EU-Commerce Directive – which holds service providers responsible for managing the content of their platform – has shown some success since, as mentioned previously, the Terrorist Trend Report of 2019 claimed there was an online decline in propaganda from jihadist organizations after the directive was implemented (Softness, 2016 & Europol, 2019). By amending the EU-Commerce Directive and EU Code of Conduct, the state began requiring media outlets to report and block content that promotes Xenophobia, racism, and violence. If companies failed to identify such content and remove it from their platforms, each corporation would face significant fines or other repercussions. The EU has claimed that its heightened regulations helped significantly to reduce the overall online presence of extremism, and the evidence from the trend report partially supports these claims. However, similarly to the criticisms of counter-narratives, it is difficult to empirically measure the exact influence that the Commerce Directive and Code of Conduct had on extremism. Furthermore, after the New Zealand shooting, the country of Australia also created legislation that now requires media platforms to remove extremist content at a quick pace or face financial consequences and – in some instances – prison time (Conifer, 2019).

Since the continent just began implementing and enforcing its regulation towards the end of 2019, it is too early to analyze its full efficacy; however, pressuring media platforms and putting companies' revenues at risk could be argued as strong motivators to keep extremist content off the Web.

Pragmatically, media accountability also has some benefits. It is impractical to expect media platforms to remove all extremism from the Internet. But, by adjusting the Communications Decency Act in the U.S. to include terrorist-specific content, and by placing incentives for companies to better manage their platforms, the overall results should show a decrease of extremist content from the general Internet. Amending the Decency Act is a rational and attainable method to hold the media accountable for radicalism and practically implement methods for identifying extremist content. Of course, as discussed earlier in this thesis, the Dark Web would be likely to gain extremists who were removed from media platforms; however, if the goal of media accountability would be to decrease the amount of radical content that everyday Internet users are exposed to, then media accountability would be fulfilling its purpose.

VI. Summary

Censorship, counter-narrative programs, and media accountability all bear benefits and limitations in both normative and pragmatic aspects. Censorship efforts would be the most controversial policy within the United States, and pragmatically challenging due to the breadth of platform capabilities that Internet users have available to them. Though counter-narrative programs would be ethically and legally beneficial to society, their difficulties to be scientifically measured give legislators legitimate reason for concern when considering such programs to effectively address online extremism. Media accountability would hold legal challenges due to the precedent of protection of corporations and the differing laws and values that make it difficult to develop specific parameters internationally; however, if media accountability were intended to make the general Internet safer for users, perhaps it could show positive results. Upon the evaluation of each option presented, this thesis concludes that, despite its limitations, the best policy for mitigating online extremism would be media accountability. Adjustments to include specific extremist content in the Communications Decency Act would be crucial to making this policy less constitutionally controversial in the United States. This amendment would hold the media accountable for extremism while upholding protections from other liability measures against the media. Overall, for any policy to be effective in combating online extremism, the goal must be to make the Internet safer for its users rather than emphasizing the impractical task of removing all extremism from the Web. If the government is to be successful in decreasing online extremism, a greater emphasis must be consistently placed on the dangers of radicalism through the use of ever-evolving technology.

References

Amble, J. (2012). Combating Terrorism in the New Media Environment. *Studies in Conflict & Terrorism*, 35(5), 339-353. doi:10.1080/105761 0X.2012.666819

Anderson, T. (1993). Terrorism and censoship: The media in chains. *Journal of Interntional Affairs*, 47(1), 127. Retrieved from http://web.a.ebscohost.com.proxy. library.cpp.edu/ehost/detail/detail?vid=0&sid=8070228 6-fbbb-44de-a28c-c0b2d044b557%40s dc-vsessmgr02&bdata=JkF1dGhUeXBlPWlwL HVpZCZzaX RlPWVob3N0LWxpdmUmc2N vcGU 9c2l0ZQ%3d%3d#AN=9402082507&db=aph

Awan, I. (2017). Cyber-Extremism: Isis and the Power of Social Media. *Springer Link*. 54(2), 138-149. Retrieved from https://doi.org/10.1007/s12115-017-0114-0.

Baugut, P., & Neumann, K. (2019). How Right-Wing Extremists Use and Perceive News Media. *Journalism & Mass Communication Quarterly*, 96(3), 696-720. doi:10.1177/1077699018803080

Behr, I. V., Reding, A., Edwards, C., & Gribbon, L. (2013). Radicalisation in the Digital Era. *Rand*, retrieved from https://www.rand.org/content/dam/rand/pubs/re search reports/RR400/RR453/RAND_R R453.pdf

Conifer, D. (2019). New Laws could see Social Media bosses Jailed for Failing to Combat Terrorist Content in wake of Christchurch Shooting. *ABC News*. Retrieved from https://www.abc.net.a news/2019-03-29/social-media-executives-jail-fine-ter roristmaterial-new-laws/10955208

Costello, M., Hawdon, J., Ratliff, T., & Grantham, T. (2016). Who views online extremism? Individual attributes leading to exposure. *Computers in Human Behavior*, (63), 311-320. doi:10.1016/j.chb.2016.05.033

Cross, S.(2013).Russia and Countering Violent Extremism in the Internet and Social Media: Exploring Prospects for U.S.-Russia Cooperation Beyond the "Reset". *Journal of Strategic Security*, 6(4), 1-24. doi: http://dx.doi.org/10.5038/1944-0472.6.4.1

Eidam, E. (2016). Apple, FBI at odds over San Ber nardino Shooter's Encrypted Device. *TCA Regional News*. Retrieved from http://proxy.library.cpp.edu/ login?url=https://searchproquest-com.proxy.library. cpp.edu/docview/1766195736?accountid=10357

Engel, R., Plesser, B., Connor, T., & Schuppe, J. (2016). The Americans: 15 Who Left the United States to Join ISIS. *NBC News*. Retrieved from https://www.nbc news.com/storyline/isis-uncoveredameri cans-15-who-left-united-statesjoin-isis-n573611

FBI, Allies Shut Two Big Dark Web Sites for Drugs, Guns. (2017). *Dow Jones Institutional News*. Retrieved from http://proxy.library.cpp.edu/login?url=https:// search-proquestcom.proxy.library.cpp.edu/ docview/1921325242?accountid=10357

Fisher, R., Cook, I., & Shirkey, E. (1994). Correlates of Support for Censorship of Sexual, Sexually Violent, and Violent Media. *The Journal of Sex Research*, 31(3), 229-240. doi: 10.1080/00224499409551756

Flynn, M. (2019). No one who watched New Zealand shooter's video live reported it to Facebook, company says. *The Washington Post*. Retrieved from https:// www.washingtonpost.com/nation/2019/03/19/ new-zealand-mosque-shootersfacebook-live-stream- was-viewed-thousands-times-before-being-removed/

Frischlich, L., Rieger, D., Morten, A., Bente, G. (2018). The Power of a Good Story: Narrative Persuasion in Extremist Propaganda and Videos against Violent Ex tremism. *International Journal of Conflict and Vio lence*, (12), 1-16. doi:10.4119/UNIBI/ijcv.644

Gross, T. (2019). For Facebook Content Modera tors, Traumatizing Material is a Job Hazard. *NPR.* [Audio Podcast]. Retrieved from https://www.npr. org/2019/07/01/737498507/forfacebook-con tent-moderators-traumatizing-material-is-a-job-haz ard

Harwood, E. (2019). Terrorism and the Dig ital Right-Wing. Contexts, 18(3), 60-62. doi:10.1177/1536504219864961

Hayden, M. (2019a). Far-Right Extremists are Calling for Terrorism on the Messaging App Telegram. *Southern Poverty Law* Center. Retrieved from https:// www.splcenter.org/hatewatch/2019/06/27/ far-right-extremists-are-callingterrorism-messag ing-app-telegram

Hayden, M. (2019b). White nationalists, who have em ployed terroristic rhetoric with increased enthusiasm in recent months, expressed solidarity with the man who police say killed at least 20 people in El Paso, Texas on Saturday. *Southern Poverty Law* Cen ter. Retrieved from https://www.splcenter.org/hate watch/2019/08/04/white-nationalists-praise-el-pa soattack-and -mock-dead

Hirsh, M. (2016). Inside the FBI's Secret Muslim Network. Politico Magazine. Retrieved from https://www. politico.com/magazine/story/2016/03/fbi-mus lim-outreach-terrorism-213765

Huey, L., (2015). This is Not Your Mother's Terrorism: Social Media, Online Radicalization and the Practice of Political Jamming. Contemporary Voices: St An drews Journal of *International Relations*, 6(2). doi: http://doi.org/10.15664/jtr.1159

ISIS Online: U.S. Rights and Responsibilities. (2017). *Counter Extremism Project*. Retrieved from https:// www.counterextremism.com/sites/default/files/ ISIS%20Online_020517.pdf

Jensen, M., Patrick, J., LaFree, G., Safer-Lichtenstein, A., Yates, E., (2018). The Use of Social Media by United States Extremists. *START,* College Park Maryland. Retrieved from https://www.start.umd.edu/publica tion/use-social-media-united-states-extremists

Killion, V. L. (2019). The First Amendment: Catego

ries of Speech. *Federation of American Scientists*, Retrieved from https://fas.org/sgp/crs/misc/IF11072.pdf

Krasenberg, J. (2019). EU Policy: Preventing the Dissemination of Terrorist Content Online. Program on Extremism: *The George Washington University*, 1-14. Retrieved from https://extremism.gwu.edu/sites/g/files/zaxdzs2191/f/EU%20Policy%20-%20Preventing%20the%20Dissemination%20of%20Terrorist%20Content%20Online.pdf

Kutner, S. (2016). The Call for Component Analyses of the Saudi Arabian Risk Reduction Initiative: An Examination of Religious Re-education's Role in the Deradicalization and Disengagement Process. *Journal for Deradicalization*, (7), 107-123. Retrieved from http://journals.sfu.ca/jd/index.php/jd/article/view/59/50

Logan M., & Richard, F. (2017). Voices Against Extremism: A Case Study of A CommunityBased CVE-Counter-Narrative Campaign. *Journal for Deradicalization*, (10),147-174. Retrieved from http://journals.sfu.ca/jd/index.php/jd/article/view/86

Lowe, D. (2017). Prevent Strategies: The Problems Associated in Defining Extremism: The Case of the United Kingdom. *Studies in Conflict & Terrorism*, 40(11), 917-933. doi: 10.1080/1057610X.2016.1253941

Mcnicol, S. (2016). Responding to Concerns About Online Radicalization in U.K. Schools Through a Radicalization Critical Digital Literacy Approach. *Computers in the Schools*, 33(4), 227-238. doi:10.1080/07380569.2016.1246883

Morgans, M. (2017). Freedom of Speech, The War on Terror, and What's YouTube Got to Do with it: American Censorship During Times of Military Conflict. *Federal Communications Law Journal*, 69(2), 145-102. Retrieved from https://search-proquestcom.proxy.library.cpp.edu/docview/2026397638?accountid=10357&rfr _id=info%3Axri %2Fsid%3Aprimo

Nele, S. & Antoinette, V. (2017). Understanding How and Why Young People Enter Radical or Violent Extremist Groups. *International Journal of Conflict and Violence*, 11(1), 1-17. doi:10.4119/UNIBI/ijcv.473

Neumann, P. (2013). Options and Strategies for Countering Online Radicalization in the United States. *Studies in Conflict & Terrorism*, 36(6), 431-459. doi: 10.1080/1057610X.2013.784568

Romm, T. & Harwell, D. (2019). Facebook, Google and Twitter face fresh heat from Congress on harmful online content. *The Washington Post*. Retrieved from https://www.washingtonpost.com/technology/2019/09/18/facebook-google-twitter-facefresh- heat-congress-harmful-online-content/

Smith, A. (2018). Public Attitudes Toward Technology Companies. *Pew Research Center*, Retrieved from https://www.pewresearch.org/internet/2018/06/28/public-attitudestoward-technology-companies/

Softness, N. (2016). Terrorist Communications: Are Facebook, Twitter, and Google Responsible for the Islamic State's Actions? *Journal of International Affairs*, 70(1), 201-215,12. Retrieved from https://search-proquestcom.proxy.library.cpp.edu/docview/1855796787?accountid=10357&rfr_id=info%3Axri %2Fsid%3Aprimo

Stieb, M. (2019). Report: Domestic Terrorism is Still a Greater Threat Than Islamic Extremism. *Intelligencer*. Retrieved from https://nymag.com/intelligencer/2019/03/domestic-terrorstill-greater-threat-than-islamic-extremism.html

Tate, E. B. (2019). "Maybe someone dies": The dilemma of domestic terrorism and internet edge provider liability. *Boston College Law School Boston College Law Review*, 60(6), 17311769,1731A. Retrieved from http://proxy.library.cpp.edu/login?url=https://searchproquest-com.proxy.library.cpp.edu/docview/2288645684?accountid=10357

Terrorism and Extremism. (2020). *Educate. Against. Hate*. Retrieved from https://educateagainsthate.com/define-extremism-terrorism-uk-2/

TerrorismSituation and Trend Report 2019 (TE-SAT). (2019). *Europol*, Retrieved from https://www.europol.europa.eu/activities-services/main-reports/terrorism-situation-andtrend-report-2019-te-sat

Tsesis, A. (2017). Social Media Accountability for Terrorist Propaganda. *Fordham Law Review*, 86 (2), 1-28. Retrieved from https://ir.lawnet.fordham.edu/flr/vol86/iss2/12/

Ubayasiri, K. (2014). The Efficacy of Censorship as a Response to Terrorism. *In the Handbook of Media and Mass Communication Theory*, 798-818. Hoboken, NJ, USA: John Wiley & Sons. doi: 798-818 doi:10.1002/9781118591178.ch43

Weekend Read: A Horrifying Pattern of White Supremacist Attacks. (2019). *Southern Poverty Law Center*, Retrieved from https://www.splcenter.org/news/2019/03/16/weekend-readhorrifying-pattern-white-supremacist-attacks

Weimann, G. (2016). Going Dark: Terrorism on the Dark Web. *Studies in Conflict & Terrorism*,

39(3), 195-206. doi:10.1080/1057610X.2015.1119546

Wilner, A. & Rigato, B. (2017). The 60 Days of PVE Campaign: Lessons on Organizing an Online, Peer-to-Peer, Counter-radicalization Program. *Journal for Deradicalization*, (12) 227-268. Retrieved from https://doaj.org/article/7b8d026d67e64cd782990d2093f7b721

Wu, P. (2015). Impossible to Regulate? Social Media, Terrorists, and the Role for the U.N. *Chicago Journal of International Law*, 16(1), 281-311. Retrieved from http://proxy.library.cpp.edu/login?url=https://search-proquestcom.proxy.library.cpp.edu / docview/1696228641?accountid=10357

47 U.S. Code 230. Protection for private blocking and screening of offensive material. Retrieved from https://www.law.cornell.edu/uscode/text/47/230