Program on Extremism
THE GEORGE WASHINGTON UNIVERSITY

THE INTERNET POLICE

This paper, part of the Legal Perspectives on Tech Series, was commissioned in conjunction with the Congressional Counterterrorism Caucus

JEFF BREINHOLT

SEPTEMBER 2019

About the Program on Extremism

The Program on Extremism at George Washington University provides analysis on issues related to violent and non-violent extremism. The Program spearheads innovative and thoughtful academic inquiry, producing empirical work that strengthens extremism research as a distinct field of study. The Program aims to develop pragmatic policy solutions that resonate with policymakers, civic leaders, and the general public.

About the Author

Jeff Breinholt is an Adjunct Professor, George Washington University Law School.

The views expressed in this paper are solely those of the author, and not necessarily those of the Program on Extremism or the George Washington University.

It is probably just a matter of time before people get in trouble solely for their social media posts. Imagine this scenario:

> *An American who, like many people, is a frequent user of social media, gets involved one night in a heated political discussion with someone she does not know. Things get heated in the back-and-forth over the course of several hours. Finally, the American gets tired, posts her last word, and retires from the exchange.*
>
> *The next day, this American tries to check her computer and finds that she no longer has an account at her social media site. She checks her email and finds a notice from the site, describing how her profile has been eliminated from the site due to a violation of the site's standards of conduct the night before. Later that afternoon, she gets a phone call from the FBI requesting that she be interviewed by two young special agents investigating her for federal crime.*
>
> *This American's political discussion has resulted in a double whammy: she has suddenly locked out of social media, and now she has to deal with possible criminal liability for her online posts.*

Is this scenario plausible?

Consider first the pressures that are felt by the social media companies. Much of American society today believes that entities like Facebook and Twitter need to prevent their services from being used by criminals. The Russia election interference is in the minds of many people, and Silicon Valley is beginning to be sued by victims who claim that their loved ones were killed in part by communication platforms that were exploited by terrorists. The remedy for this pressure is to bolster the tech companies' compliance departments, and to incentivize them to write algorithms to minimize their risk of exploitation of their platforms by criminals. I believe these companies could write code that constantly monitors their customers' online communications. The algorithm could constantly identify the name of the customer and exactly what she posted, and why the communications might require action by the company. A set of humans could be

responsible for determining whether what the server identifies is actionable. If it is, the humans make a unilateral decision to de-platform the customer. A form email to the customer follows.

What about the FBI? Americans expect the FBI to keep them safe from terrorism. When the FBI fails, it becomes a scandal. Aggressive Congressional oversight is initiated, public hearings held, and the Inspector General gets involved. There may even be an Independent Commission. All of these entities want to determine how the FBI missed the warning signals, so the it can be reformed and such catastrophic errors avoided in the future.

For terrorism at least, we seem to be moving towards a consensus that, in order for the FBI to effectively do its job, it needs the cooperation of social media companies. What about crime in general? U.S. law since 1970 has required American financial institutions to report suspicions that their customers are engaged in crime. Where they fail to do this, they get into trouble with their regulators.

Currently, social media companies are not federally regulated, but it is possible this could change. When they are, might they be required to “know their customers” (like banks are now) and notify law enforcement of possible criminal conduct by them?

The factor that makes this scenario so plausible is that Silicon Valley will be the first to see crimes that can be committed exclusively on the communication platforms they offer to the public. The platforms will occasionally generate troublesome communications, which tech company compliance officers might consider in taking some corporate action. In some cases, there will be no action. For others, it might just be a matter of cutting the customer off under the terms of service, or some other more minor form of discipline. In more extreme cases, they might refer to matter to the FBI for further proactive action. For the third scenario, the company and all such companies offering similar services) will need to know exactly what the trigger point is for FBI involvement.

This question intrigued me as I thought about this future vision. If American law enforcement and the tech companies get closer due to their commonality of interests,

what might the police tell Silicon Valley about what type of online communications should be referred to them for action?

This article focuses on the most extreme cases – where an individual commits a federal crime exclusively by typing a message into a computer or cell phone. It is not focusing on whether an online post might be a single overt act in a wide-ranging criminal conspiracy, since the pertinence of that post would require some familiarity with what the FBI knows about the scheme, which the FBI would not disclose. The tech companies, as good as they are, will not be able to program their system to uncover such non-obvious criminal posts.

Are Americans ever prosecuted solely for their online communications? The answer is yes, because of the enforcement of two federal criminal statutes.

First, there is 18 U.S.C. § 875, which criminalizes the act of transmitting in interstate commerce any threat to injure the person of another. The second is 18 U.S.C. § 2261A, which criminalizes the use of an interactive computer service to engage in a course of conduct that causes substantial emotional distress. Each of these statutes has, in recent years, been used to prosecute people for their online activity. In many of these cases, indicted defendants argued that the prosecution violated their First Amendment rights. How courts have handled these claims point to factors that the FBI should look to in deciding whether a crime has been committed or whether the customer was merely engaged in heated rhetoric. These cases also suggest what the FBI could tell social media companies about what they should look for.

Here's how it works in court:

When a person is charged with a crime for their online communications, they frequently claim that the prosecution violates their First Amendment rights. They do this in different ways. Sometimes, they claim the statute they are charged with violating is unconstitutional because it criminalizes free speech. This is referred to as a "facial challenge" to the statute based on overbreadth. In other cases, they claim that the statute is being applied to their conduct in an unconstitutional way, because their

specific alleged activity is constitutionally-protected. This is referred to as an "as applied" challenge.

The ability of criminal defendants to make these arguments was impacted by the Supreme Court's decision in *Elonis v. United States*, 135 S. Ct. 2001 (2015). There, the defendant was charged and convicted of violating § 875, based on comments he posted on Facebook. Specifically, he posted self-styled rap lyrics containing graphically violent language and imagery concerning his wife, co-workers, a kindergarten class, and state and federal law enforcement. These posts were often interspersed with disclaimers that the lyrics were "fictitious" and not intended to depict real persons, and with statements that Elonis was exercising his First Amendment rights. Many who knew him saw his posts as threatening, however, including his boss, who fired him for threatening co-workers, and his wife, who sought and was granted a state court protection-from-abuse order against him. When Elonis's former employer informed the FBI of the posts, it began watching his Facebook activity and eventually arrested him for violating § 875.

At his trial, the court instructed the jury that it could only find Elonis guilty if a reasonable person would foresee that his Facebook posts would be interpreted as a threat. Elonis had argued that the instruction should have required the jury to find that he had subjective intent to make a "true threat," irrespective of how the posts were received. The jury instruction controversy went up on appeal after Elonis' conviction, and the Third Circuit affirmed the court's decision. Elonis then sought review by the Supreme Court.

The majority opinion, authored by Chief Justice Roberts, noted that the text of § 875 did not contain a mens rea requirement. This left the court to determine what the proper intent standard was for conviction. It ultimately decided that § 875 convictions require the prosecution to prove that the accused's communication was for the purpose of issuing a threat or with knowledge that the communication will be viewed as a threat. Anything less – like judging the communication by the impact it would have on a "reasonable person," as in Elonis' case - would mean that people could be prosecuted for negligence, which was unconstitutional.

The immediate effect of *Elonis* was an effort by some § 875 convicts to overturn their convictions. See *Shah v. United States*, 2016 WL 6762748 (S.D.W.V. 2016); *U.S. v. Sherbow*, 2016 WL 1272907 (D.C. Cir. 2016). It also meant that, going forward, the government could only prosecute people for whom there is evidence that they had a subjective intent to make a threat. Threats cannot now be judged on the objective "reasonable person" standard. Merely negligent online posters cannot be prosecuted.

The *Elonis* decision, while a victory for the defense, hurt the ability of future § 875 defendants to claim that the statute was facially overbroad, since the Court had essentially limited § 875 to defendants who were accused of making "true threats" that were not protected by the First Amendment. Overbreadth challenges, after all, require the claimant to show that the statute has a "substantial" impact on First Amendment-protected activity. Post-*Elonis* defendants now have a difficult time explaining how § 875 could potentially be used against people for exercising their free speech rights. The Court basically cabined the reach of the statute, making facial overbreadth challenges very difficult.

This is not to say that the First Amendment is never in play in these cases, for defendants can still claim that their particular prosecution infringes on their constitutional rights, which they do in "as applied" constitutional challenges. The problem is that the question of the defendant's intent is generally a question for the jury, which means that trial courts generally do not dismiss these indictments before trial.

Still, in my opinion, prosecutors never want to lose a case, and most prosecutors will not bring prosecutions where there is not some evidence that the defendant actually intended to engage in criminal activity. This might be small comfort to civil libertarians, who are sometimes distrustful of explanations that rely on the concept of prosecutorial discretion.

With regard to the interstate stalking crime, 18 U.S.C. § 2261A, the elements are explicit in terms of what the statute requires on the defendant's intent, and there has so far been no need for the Supreme Court to clarify or cabin the statute to avoid unconstitutionality. Still, § 2261A defendants frequently claim that either the statute is

unconstitutional or that its application to their conduct infringes on their First Amendment rights.

If the question is what the FBI and prosecutors would look for from social media companies regarding their customer's online activity, the best guidance comes from distilling what courts have said about § 875 and § 2261A prosecutions when they are constitutionally challenged. After all, in my opinion, American law enforcement is not in the habit of trampling on people's rights by initiating bad prosecutions.

Let's look first at the few cases where courts have credited defendants' claims that they are being punished for constitutionally-protected online communications.

The only § 2261A online communication prosecution so far to be dismissed on First Amendment grounds was *United States v. Cassidy*, 814 F.Supp.2d 574 (D.Md.2011). William Lawrence Cassidy had joined the leader's religious community and then been cast out, was charged for a harassing campaign of Twitter messages directed at a religious leader. He created a number of Twitter profiles, and used those profiles and multiple blogs to direct thousands of derogatory messages to the leader. The district court held that the defendant's speech was protected expression, because despite their bad taste they challenged the target's "character and qualifications as a religious leader." Of particular importance, the court concluded that the "Indictment amounts to a content-based restriction because it limits speech on the basis of whether that speech is emotionally distressing to [the leader]."

Several of defendants thereafter tried to fit their prosecutions into the *Cassidy* precedent.

Jovica Petrovic was unsuccessful in getting his ex-girlfriend to take him back. He told her he had sexual pictures of her that he would publicize if she did not agree. When she refused, he set up a website that has the word "slut" in the URL. Petrovic reported his site was "huge," containing "20,000 or 30,000 pages" of material reflecting months of preparation by him. The website contained links to dozens of images of the girlfriend posing in the nude or engaging in sex acts with Petrovic. Visitors to the site could view

scores of pictures of her children and other family members by clicking on a link next to the pornographic material. Several photographs of her performing a sex act with Petrovic were repeatedly and prominently displayed throughout the website, including on the site's home page. Petrovic also posted thousands of pages of the text messages she had sent him. The messages were color-coded by speaker and organized chronologically, with the most private and embarrassing messages given special pages to increase readership. Petrovic posted the pictures of the blood from the woman's suicide attempt, further highlighting her suicidal thoughts and history. Private information about her and her family was also revealed, including her contact information and the social security numbers of her children. After learning of the website, the girlfriend "had a breakdown" and "wanted to die."

Petrovic was convicted under § 875(d) and § 2261A. On appeal, the Eighth Circuit had no trouble affirming the constitutionality of his convictions. First, it found that Petrovic's communications were integral to this criminal conduct as they constituted the means of carrying out his extortionate threats, and were therefore not protected by the First Amendment. Second, the court concluded the posts where matters of purely private significance, where the First Amendment protections are often less rigorous, because restricting speech on purely private matters does not implicate the same constitutional concerns as limiting speech on matters of public interest. It also noted that the victim was not a public figure and Petrovic's posts revealed intensely private information about her, and that the public has no legitimate interest in the private sexual activities of her or in the embarrassing facts revealed about her life, which distinguished the situation from that in *Cassidy*. *U.S. v. Petrovic*, 701 F.3d 849 (8th Cir. 2012).

David Thomas Matusiewicz had an even worse case, in part because the victim, his ex-wife, was murdered by his brother in the course of Matusiewicz' harassment of her in a custody dispute. Although Matusiewicz was not charged with the murder, he was charged under § 2261A. Prosecutors planned to present evidence that Matusiewicz and his family posted accusations against the ex-wife online, sending accusations against her to the school that one of the children attended and her church, and soliciting their

friends' assistance in visiting her home to monitor her. The court rejected his First Amendment-based motion to dismiss, largely on the basis of the distinction between the private communications at issue and the public issues in *Cassidy*. *United States v. Matusiewicz,* 84 F.Supp.3d 363 (D. Del. 2015).

Shawn Sayer, seeking to harass a former partner, posted an online ad on Craigslist, created fake Facebook and MySpace accounts, and posted explicit photographs of her on pornography websites. In these postings, he impersonated her and invited men to her house for sexual encounters, leading a number of men to appear at her door. His § 2261A conviction was affirmed over his First Amendment objections, in part because Sayer could not articulate how his online communications were protected by the First Amendment. *U.S. v. Sayer*, 748 F.3d 425 (1st Cir.2014).

United States v. Osinger, 753 F.3d 939 (9th Cir.2014) involved similar factual circumstances. Christopher Osinger, repeatedly contacted his ex-partner asking her to restore their relationship. After being refused, Osinger created a fake Facebook page in his ex-partner's name which included sexually explicit photographs of her. Osinger also sent explicit pictures of his ex-partner to her current and former co-workers. A jury convicted Osinger under § 2261A. The Ninth Circuit rejected an as-applied First Amendment challenge to the prosecution, also concluding that Osinger's speech was not protected expression: "Any expressive aspects of Osinger's speech were not protected under the First Amendment because they were 'integral to criminal conduct' in intentionally harassing, intimidating or causing substantial emotional distress to [the victim.]"

Finally, Kris Sergentakis was upset with his co-worker who was involved in an investigation that resulted in Sergentakis' fraud conviction involving the charity where they both worked. After he got out of prison, Sergentakis created a series of website pages and a Facebook page in which he made a number of allegations about the co-worker's supposed animal cruelty and pedophilia. Charged under § 2261A, Sergentakis claimed in a motion to dismiss that his online activity was free speech.

The court rejected the comparison to *Cassidy*, where Sergentakis had argued that his communication involved the operations of a major, public charity, and the salary and actions of its then-Chief Financial Officer and Chief Executive Officer, and were therefore of public interest and protected by the First Amendment:

> Simply put, this prosecution concerns the defendant's campaign of personal attacks against Walter through letters, emails, and the Internet, concerning allegations of child molestation, animal cruelty, case fixing, and rape, among others. To the extent that the defendant's speech, as he contends, concerns [the charity's] operations, executive compensation, and management, those statements do not form the basis of the Indictment, and, at most, appear to be a thinly veiled attempt to immunize the defendant's personal attacks on Walter by claiming to speak on public issues.

The court found that the context in which these statements were made is particularly instructive, noting that while some of Sergentakis' posts could be protected by the First Amendment in other circumstances, these were not statements made purely for altruistic reasons as part of a critical campaign against the charity in which he had long been engaged. Instead, the posts were not protected under the First Amendment because they were "integral to criminal conduct in intentionally harassing, intimidating or causing substantial emotional distress to" the victim. *U.S. v. Sergentakis* 2015 WL 3763988 (S.D.N.Y. 2015).

In addition to *Cassidy* (which was a § 2261A case), there have been a couple of defendants who were able to successfully defend against § 875 charges focusing on their online communications, and these cases should be of interest to prosecutors and agents, and to social media companies that might refer certain troublesome communications to them.

William A. White was able to partially dodge two bullets despite himself and his bizarre online behavior. White is the "Commander" of the American National Socialist Workers' Party, which he formed in 2006. He conducted activities from his home in Roanoke, Virginia, promoting his neo-Nazi white supremacist views by publishing a white

supremacist monthly magazine; by posting articles and comments on his white supremacist website, "Overthrow.com," as well as on other similar websites, such as Vanguard News Network Forum; and by conducting a radio talk show.

White's first legal troubles started when was charged with sending threatening emails to an employee of a bank with who he was having a dispute, sending threatening letters to black tenants who had reported housing discrimination to HUD, making a threatening call to the diversity office of a university, and making threatening emails and online posts to a Canadian human rights lawyer named Richard Warman who had been fighting white supremacism. White was largely convicted, although the court did grant his motion for acquittal for the count involving the alleged threats to the Canadian lawyer. The court had rejected the motion for acquittal on the other counts. After his conviction, White appealed his convictions, while the government cross-appealed appealed the court's granting the Rule 29 motion as to Count 6 (the Warman posts).

The Fourth Circuit seemed to have no trouble affirming most of White's counts of conviction (albeit this was before *Elonis*). However, it also affirmed the trial court's granting White's motion for judgment of acquittal on Count 6, writing:

White's communications directly and indirectly to Richard Warman were part of a protracted campaign to oppose Warman's work in Canada, fighting neo-Nazi and white supremacy groups. Except for the two communications charged in Count 6, however, these communications were presented only as context, and as context, they were insufficient to elevate the communications in Count 6 to true threats.

The first of the two communications forming the basis for the conviction on Count 6 was a February 8 posting on a web-site that referenced the recent firebombing of a Canadian civil rights activist's house with the subscript, "Good. Now someone do it to Warman." The second, in March 2008, was again a posting on White's website indicating that Warman "should be drug [sic] out into the street and shot." It also asserted that "Richard Warman is an enemy, not just to the white race but of all humanity and he must be killed." These communications

> clearly called for someone to kill Richard Warman. But neither communication actually provided a threat from White that expressed an intent to kill Warman. While a direct threat of that type would not always be necessary, for White to have called on others to kill Warman when the others were not even part of White's organization, amounted more to political hyperbole of the type addressed in Watts than to a true threat. Moreover, the two communications forming the basis for Count 6 were posted to neo-Nazi websites and not sent directly to Warman.
>
> While a direction to others to kill Warman could have amounted to a threat if White had some control over those other persons or if White's violent commands in the past had predictably been carried out, none of that context exists in this case. *In short, the communications that formed the basis of Count 6 were expressions not directed to Warman but to the public generally and did not communicate an intent to take any action whatsoever. In these circumstances, we agree with the district court that the communications fell short of being true threats.... While the two communications for which White was indicted, along with the context surrounding them, may have undoubtedly frightened Warman, those communications at most conveyed a serious desire that Warman be harmed by others but did not convey a serious expression of intent to do harm from the perspective of a reasonable recipient. Accordingly, we affirm the district court's judgment of acquittal on Count 6.*

U.S. v. White, 670 F.3d 498 (4th Cir. 2012)(emphasis added.)

White had gotten somewhat lucky, but he found himself in trouble the next year, when he faced sanctions based on the content of several Internet blog postings that White authored during the course of the underlying civil HUD housing discrimination dispute. These postings paired White's criticisms of the court, its processes, and the litigants appearing before it, with expressions of his anti-Semitic and white supremacist views and, at times, with the personal, identifying contact information of the attorneys

involved in a lawsuit (to which White was not a party). The court ultimately opted not to award the sanctions, because the relevant posting by White were protected by the First Amendment. Doing so, it went through the history of the "true threat" doctrine in the Fourth Circuit, and focused on several factors: White's language, the context, and his history of communicating his views. It concluded that White's posts did not constitute "true threats" because the record failed to reveal any disruption to the underlying litigation, nor the imminent likelihood of such disruption or interference, and no clear and present danger.

In concluding that White's speech is constitutionally protected, the Court does not minimize the real fear of harm and intimidation that Mottley and his family experienced as a result of his conduct. The Court strongly disapproves of the method by which White sought to express his views in this matter. Despite its protected status, the Court finds White's conduct to be reprehensible and, again, emphasizes that minute or subtle changes to the language or context may have resulted in the exclusion of his speech from First Amendment protection. The significance of this point should not be lost on White or on any other similarly situated person in light of the Court's ultimate ruling., in our democratic society, when presented with even caustic or abusive protected speech, "we do not quash fear by increasing government power, by proscribing [our fundamental] constitutional principles, and silencing those speakers of whom the majority disapproves."

In re White, 2013 WL 5295652 (E.D. Va. 2013)

Ashton R. O'Dwyer Jr. was charged under § 875 for an email he sent to a bankruptcy judge which complained about his need for a release of some of his debtor assets so he could purchase prescription medication. The message stated:

Maybe my creditors would benefit from my suicide, but suppose I become "homicidal"? Given the recent "security breach" at 500 Poydras Street, a number of scoundrels might be at risk if I DO become homicidal. Please ask His Honor to

> consider allowing me to refill my prescription at Walgreen's, and allowing me to pay them, which is a condition for my obtaining a refill.

O'Dwyer claimed that the language was not a threat because this is merely how he speaks after Hurricane Katrina, and that the message was simply a cry for help. The court agreed, dismissing the charge:

> These e-mails place the allegedly offending e-mail in context. At no point did the Defendant threaten anyone. His e-mails, while filled with coarse language, did not threaten bodily harm. Phrases taken out of context could suggest a threat, but reading the sentences as a whole, no threat as a matter of law was made.

U.S. v. O'Dwyer 2010 WL 2606657 (E.D. La. 2010) The Fifth Circuit affirmed, noting:

> O'Dwyer made his allegedly threatening statement in an e-mail transmitted to a bankruptcy court employee, with a message for Judge Brown, in which he never identified any individual whom he intended to harm. The most he said was that "a number of scoundrels might be at risk." We conclude, based on the language of O'Dwyer's statement, and in light of his documented history of using coarse and hyperbolic language in prior court proceedings, that no reasonable jury could find that O'Dwyer's communication constitutes a true threat.

U.S. v. O'Dwyer 443 Fed.Appx. 18 (5th Cir. 2011).

These cases should be scrutinized by federal prosecutors if they ever receive real-time data from social media companies. It might be one thing for the compliance departments at Facebook and Twitter to cut a customer off from their service for abusive online posts. It is quite another to parlay this information into a cognizable criminal prosecution in American courts, where full context matters.