Mais Hriesh

PEN America

Rosenberg Award

09/13/2017

Hateful Speech and Trolling

PEN America's community of writers demands from PEN an immediate stepped up effort to equip them to respond to hateful speech and defend the open discourse upon which their crafts and livelihoods depend - and from which the broader public benefits. PEN America has documented several dozen cases on New York City based writers who have been targeted by intensifying campaigns of mostly online hateful speech based on their writings, but also tied to their racial, gender or religious identities, or their immigration status or country of origin. In some cases writers have reported departing from social media platforms and public discourse for fear of being trolled and targeted for the expression of their views.

The working title of this effort is Write Versus Wrong, will educate and galvanize the writers community with tools to combat hateful speech. PEN aims at creating and disseminating a toolkit specially designed for writers including how to understand the legal aspects and related free expression issues governing hateful speech, effective measures and resources for countering hateful speech directed at them personally, and what writers can do to respond to hateful speech and trolling, whether they are a victim, witness or ally. PEN will draw on resources and expertise from groups and individuals with legal, technological and psychological expertise relevant to the issue of hate speech and trolling online.

Laura Macomber was hired especially to run and monitor this project. Macomber claims that "online hate speech and trolling have reached a crescendo in recent years, especially in this charged political time we live in, frequently drowning out the voices of writers wishing to engage in civil online discourse and consistently targeting journalists, authors, critics, bloggers and many others publishing content on the internet. The impact of online hateful speech and trolling on an individual's productivity, professional advancement and mental health is well-documented and pervasive."

According to the Pew Research Center, 40% of adults experience online harassment in their lifetime. In some cases writers have reported self censoring or departing from social media platforms and public discourse for fear of being trolled or targeted for the expression of their views. As part of Write Versus Wrong, PEN will disseminate a preliminary survey to writers whose lives may have been impacted by online hateful speech and trolling. With the information collected from these surveys, PEN will be in a prime position to assess writers' needs when it comes to combating online hate, as well as to identify any gaps in online community standards, law enforcement, or cyber security that should be addressed.

My role at PEN was mainly research based, since the project was just starting when I got to PEN, my supervisors needed to have an overview of all the different sections and angles we might have to research and study to design and start the project. Therefore, have a clearer idea about how the toolkit should look like.

I started my very wide research for the project by collecting cases of people and writers who have been or are being trolled. I knew that there was going to be a good amount of data waiting for me to collect, but I didn't think it was going to be that generous. I created a spreadsheet to organize the cases I collected that had the following categories: Name, location, gender/gender identity, race/ethnicity, writing affiliation, description/hate speech experience, tags and contact information. After analyzing the spreadsheet it appeared that the majority of writers who experienced trolling were women who identified as feminists. Both, white and of colour. The same results appeared at The Pew Research Center, as it has found that "women are more likely than men to be tormented on the Internet, with about one in four women between the ages of 18 and 24 reporting being stalked or sexually harassed online — rates two to three times higher than among men of the same age. Men were slightly more likely than women to report online harassment, but they usually experienced less severe forms of it. In fact, while a 2014 survey found that men receive more negative messages on Twitter in general, the group that received even *more* harassment was female journalists — that is, some of the women who rely most on the Internet as a way to share their thoughts and simply do their jobs. Women of color and members of the LGBT community tend to face especially fierce online harassment."[1] The spectrum of the definition of hateful speech is very wide and it's very hard to categorize it. Therefore, the intersection between the definition of hateful speech and free speech is very problematic. This perplexity can be derived from the First Amendment.

The First Amendment guarantees the right to freedom of expression. "Hate speech" is covered by the First Amendment but may be punishable if it incites attacks or is otherwise

---

[1] Rainie, Lee, et al. "The Future of Free Speech, Trolls, Anonymity and Fake News Online." *Pew Research Center: Internet, Science & Tech*, 29 Mar. 2017,

threatening beyond insulting a person or a group on the basis of national origin, ethnicity, color, religion, gender, gender identity, sexual orientation, or disability.[2] Though it is fairly complex in certain cases to determine whether a speech is punishable or not since offensive speech is not always hateful speech.

The U.S Supreme Court has carved out some exceptions to the First Amendment. These exceptions include "Fighting words", words that provoke the listener to resort to immediate violence". The First Amendment also doesn't protect statements that are meant to incite particular listeners to take immediate illegal action and are likely to have that effect, or "Obscenity and Child Pornography" and "defamation."[3] There are some other exceptions added to this list, but they are somehow less affirmative.

After spending some time with this initial research, we moved to the next step. Which is to identify the harm that is caused to journalists and writers by hateful speech so we can understand and have a better idea of how to help them. To our surprise, not too many people wrote scientific articles about the negative psychological effects of hateful speech. That's why Macomber's answer to the question I asked her when I interviewed her "What additional scholarship would be helpful for you that you think will help you do your job better, was: "to me there seems to be a large gap in the research that's been conducted into the lives of victims of online hate speech and trolling: what is the psychological toll, the personal and professional toll?

[2] "Hate Speech and Hate Crime Update & Resources • Advancement Project California."*Advancement Project California*. N.p., 27 Jan. 2017.
[3] Gjelten, E.A. "Does the First Amendment Protect Hate Speech?" *Lawyers.com*. Lawyers.com, 18 May 2017.

If we can begin to document this more clearly then we can arm ourselves with stats and stories to help change the tide of popular opinion around hate speech online."

One piece of scholarship that was very helpful to us was "Experiencing Hate Speech: Perceptions and Responses to Anti-Semitism and Antigay Speech" by Laura Leets. In her research Leets said that: "other scholars have argued persuasively that words (van Dijk, 1987) and pornographic pictures (MacKinnon, 1993) may hurt people in deep and permanent ways. Along these lines, people have reported that psychological or emotional pain can be a worse experience to bear than physical pain or material loss (Leets & Giles, 1997. 344)."

In her research, Leets emphasizes on the fact that certain words might be a trigger certain people and be of no effect on others. Anne Walsh from the No Hate Speech Movement Ireland had similar statements. She said: "One of the largest consequences of hate speech is the effect that it can have on a victim's mental health. Many of those who are victims of Hate Speech may self harm, or become suicidal, their confidence and self esteem may be affected, as well as many more consequences. What to one person is a simple word, to another is much worse. While many people believe that how someone interprets their words is not their problem, this is not true. It is your responsibility to ensure that your words aren't hateful."[4]

A handful of writers came up with some strategies to face online trolling. Dylan Marron, one of the most known names in this field is working on a podcast called "Conversations with People Who Hate Me". The podcast takes hateful conversations online and turns them into

---

[4] *No Hate Speech Movement Ireland*, www.nohatespeech.ie/effects-hate-speech-lorna-costelloe.

productive conversations offline. a podcast where he has extended conversations with some of the folks who have said the most negative or hateful things about him on the internet. It was selected as a "podcast pick" by USA Today & The Guardian.[5] Although there are some people trying to come up with strategies to shut and face the trolls, the most popular solution stated on the internet to deal with trolls is to "not feed the trolls". On the one hand, it might be better to ignore the trolls to avoid the negativity and the issue escalating. On the other hand, why should those who are being trolled be silenced when it's their right to speak up and defend themselves?

Bethany Mandel, a conservative essayist at the Federalist who has been awash in a deluge of social media abuse and antisemitic trolling since she began denouncing Donald Trump, finally purchased a handgun after being virtually harassed. Mandel is an aggressive online presence who often inveighs against Trump and his "outrageous Jewish hypocrisy."[6] To Mandel's estimation, up to 1000 Twitter trolls by Trump supporters were taunting her due to her views and opinions on Trump. She was verbally harassed and received straight up death threats. This included "you deserve the oven," complete with the image of a Domino's Pizza oven—this from an apparent Trump fan who goes by the Twitter handle @dinguscout. Therefore she decided to purchase a handgun, even though her husband and family didn't like that.

Amanda Hess is another female writer who fell a victim of a Twitter Account that was set up with the sole purpose of harassing her and making death and rape threats. Hess is an

---

[5] "ABOUT." *Dylan Marron*, www.dylanmarron.com/about/.

[6] Bonazzo, John. "Jewish Journalists Warn Anti-Semitic Trolling 'Not Going Away' After Election." *Observer*, Observer Media, 15 Dec. 2016

American journalist who has written for magazines including Wired, ESPN, and Elle. She writes about internet culture for the Arts section and contributes regularly to the New York Times Magazine. Hess says she was used to receiving trolls as a female writer who writes about sex amongst other things. But this specific troller was different. He took it to the next level and set up a twitter account just in the purpose of harassing Hess. He used the username "headlessfemalepig" and started by sending a tweet about Hess' physical appearance. "I see you are physically not very attractive. Figured," then he continued: "I am 36 years old, I did 12 years for 'manslaughter', I killed a woman, like you, who decided to make fun of guys cocks." And then: "Happy to say we live in the same state. I'm looking you up, and when I find you, I'm going to rape you and remove your head." There was more, but the final tweet summed it up: "You are going to die and I am the one who is going to kill you. I promise you this." Hess was confused about how she should feel because on the one hand she thought that it seemed unlikely that she was going to be killed by this serial murderer. On the other hand, she couldn't hide her feelings of fear. She called the police and they showed up two hours later. After they've done their investigation they told her that "this guy could be sitting in a basement in Nebraska for all we know." The fact that the stalker had said that he lived in the same state as Hess, and had plans to chase her to her house was dismissed as just another online scheme. Hess writes: "no matter how hard we attempt to ignore it, this type of gendered harassment—and the sheer volume of it—has severe implications for women's status on the Internet. Threats of rape, death, and stalking can overpower our emotional bandwidth, take up our time, and cost us money through legal fees, online protection services, and missed wages. I've spent countless hours over the past four years logging the online activity of one particularly committed cyberstalker, just in case. And as the Internet becomes increasingly central to the human experience, the ability of women

to live and work freely online will be shaped, and too often limited, by the technology companies that host these threats, the constellation of local and federal law enforcement officers who investigate them, and the popular commentators who dismiss them—all arenas that remain dominated by men, many of whom have little personal understanding of what women face online every day."[7]

Almost all the staff that I interviewed at PEN believed that there is a direct relationship between hatred and ignorance. They believe that people come to hate another group of people because of their prejudices and misinformation about this specific group of people. They try to reduce this notion by putting all different groups of people on the social arena. They do that through literature and through the power of the word. By ignorance, they don't mean that all the people that "hate" or condemn a specific group of people are completely ignorant and unknowledgeable. These people can be very knowledgeable and think that they've done their homework in the best way possible to prepare their argumentation of why they hate a certain group and why it should be hated. The argument that the staff at PEN endorses is that these people go into a process of "cherry picking" and pick the negative things they find or see in the "hated group" and leave out the somewhat positively controversial things aside and avoid to discuss them or refer to them since they don't fit into, or invalid to their argumentation. Laura Macomber, my supervisor, had another view on this. In her interview she said: "Even after all the work I've done, all the stories I've covered in which people are victimized or physically harmed by another person or group's hatred, to this day I remain shocked by the hateful things people continue to do and say to one another. As a means of processing hate, and trying to assess

[7] Hess, Amanda. "The Next Civil Rights Issue: Why Women Aren't Welcome on the Internet."*Pacific Standard*, 6 Jan. 2014

the damage it does to people, I've recently taken a more anthropological/evolutionary approach to understanding it. Being able to hate another human or group was once a very useful survival tool for human beings. Groups and clans were competing for resources, and in order to secure those resources for their own survival, it was necessary to dehumanize the competing group; hatred leads to dehumanization, and dehumanizing the 'other' is a means of absolving oneself of the crimes you feel you must commit against one's fellow humans. Of course today, to look at the world's stable democracies, hate and the violence it breeds begin to feel like a wholly outdated system for disagreeing groups and individuals to resolve their issues with each other. When I think about hate, and what spawns it, I assume it falls to those with few resources for processing their troubles and disappointments. But of course, there are plenty of people who have been given every opportunity, who have very little to be disappointed in, and yet they still find it within themselves to hate. I think hate is something we all have a capacity for and we all need to work on eradicating."

When I asked the question of "what additional research and material you would like to see to help you understand the issue of hatred and do your work better?" four out of the five people I interviewed answered that they would like to study and see "what changes people's minds" or "how can they be changed". They talked about the difficulty to change people's views and opinions. When people grow up taking some ideas and matters by default it is very hard to change them. Even if you show statistics and well researched facts that contradicts their point of view, it is still not guaranteed that people would shift or change their opinions.

One of the useful ways they thought of is a report about children's literature. The report examines two issues. The first issue was that often, in certain books that are assigned in a school classroom, parents or community members will challenge the books that are assigned because they are worried about what the books have to say. Therefore, often it is books that have LGBTQ or other diverse characters that are most likely to be challenged, and parents would tell the schools that they shouldn't be making their children read such books that carry "unconventional" characters and messages. PEN focuses on the fact that a disproportionate number of the books that get challenged in that way, feature diverse characters. The second issue is the problem that in children's literature, there is not nearly the representation of characters of color or the LGBTQ identities to begin with. Most children books feature white straight children. PEN believes that it really matters to see yourself in books as you're growing up. It's the type of thing that if you are a white reader that may have never even occurred to you, because you never experienced this lack of representation. In this project, PEN tries to raise awareness and acceptance of diverse characters in children's literature, so these children grow up to be tolerant and aware humans with no prejudices to those who are different.

James Tager is the Program Manager for PEN America's Free Expression Program. He stated: "I think that civic education is the way to go. Ideologies of hatred are obnoxious ideologies, and it deserves no place in the political sphere. As a neutral statement I can say that they shouldn't exist. However, as a human being and a citizen in a democracy I recognize that it's not our place to mandate that people don't believe in certain things. Not only that this is impossible but also I, as a human being, human rights advocate and as an American citizen, believe to be objectionable. The government can't say "don't believe in Nazis or we'll arrest

you.' Also it's kind of abductive, because extremist people might use the theory of victimhood to protect themselves. And we've seen that very much recently from the alt right movements and Trump supporters. Therefore, it is impossible, unethical and impractical to legislate any type of lack of ideology, even if it's a hateful ideology. So, that leaves us persuasion. How do you persuade people that hatred is bad? And I think the best way to do that is through civic education. Now, there is also cultural communization and through it, exposing people to the object of their hatred. I'm sure that many people in this country who hate Muslims and are Islamophobic have never met or had a conversation with a Muslim their entire life. So that's the place of the community groups and the civil society to increase and foster the connections between people. I believe it's a way that can break hatred."

There wasn't a specific source of scholarship that the staff at PEN was regularly reading to rely on and inform their views on the issue of hatred. But they all said that they try to read as many diverse sources as they can to have a more clear picture on where people are standing. This includes Breitbart. Laura Macomber said: "I come from a journalism and storytelling background, so a lot of the research I've done into hate has come from stories shared by individuals: for example, people who have 'recovered' from their hate. I take a human approach to understanding hate, I like to try to understand it through a human's capacity for change, for rejecting hatred - i.e. a gang member who went to prison for twenty years only to be transformed into a pacifist; or a former white supremacist who has come to terms with his own racism. I seek excellent journalism about these subjects - long form pieces and documentaries that touch on hate. These stories may be few and far between and don't necessarily represent the larger structural inequalities that engender hatred in most stable democracies. But once again, this is

where an evolutionary/anthropological approach to understanding hate has been helpful - if I can understand why it exists to begin with, I can begin to understand how we might nurture its reduction."

Some people see the internet as an anarchical space where those who operate online must be thick-skinned enough to protect themselves from verbal attacks in the name of free speech protection. Danielle Keats Citron, rejects this view. Danielle Citron is a well known expert on cyber harassment that focuses in her work mainly on information privacy law, cyber law, administrative law, and civil rights. In May of 2016, she released a book called "Hate Crimes in Cyberspace". Citron sees that cyber-harassment is a matter of civil rights law, and legal precedents as well as social norms of decency and civility must be leveraged to stop it.

Laura Macomber writes after her interview with Dr. Danielle Citron: "First, Citron is in favor of amending Section 230 of the Communications Decency Act. In a nutshell, Section 230 protects website owners from being held directly liable for content third parties post on their site, unless federal criminal law or intellectual property is involved. That provision is, for most modern media entities, a lifesaver. It means they're not liable for the content people post in the comments. However, it's also been a lifesaver for revenge porn titans like Hunter Moore. Moore only got busted when the federals decided he'd violated a federal criminal hacking law. Citron would like to see the provision amended to "exclude the very worst actors: sites that encourage cyber stalking or nonconsensual pornography and make money from its removal *or* that principally host cyber stalking or nonconsensual pornography." Second, Citron is a believer in actively legislating against the worst bits of the internet. She has, for example, been one of the

people lobbying for <u>properly-drafted</u> anti-revenge porn laws. Those laws, which are showing up all over the country, criminalize the posting of someone's nudes without their consent."

In addition to criminal laws, Citron also proposes that we develop a scheme of civil rights remedies for victims of internet harassment, mostly by amending current state civil rights laws and federal laws like Title VII of the Civil Rights Act of 1964. Citron argues that such legislation would not fall afoul of the First Amendment because the First Amendment has never protected threats or harassment anyway.

My work at PEN has been eye opening, and it allowed me to explore the notion of free expression more in depth and how problematic it can be in some cases, especially in the era we live in right now. Hatred in general is very hard to combat. But when it comes to online trolling, tracking it becomes even harder, since most trollers are anonymous. It is very problematic for social media platforms to come up with a concrete solution to banish trollers without contradicting their free speech policy. The only viable solution for now is to have a massive civil society and grassroots awareness campaigns to inform people how to be responsible online. Meanwhile, PEN America, in partnership with other organizations, is working on helping writers and active internet users with providing them with the toolkit.