

How online Islamophobia goes unmonitored and the toll it takes in reality.

Identity and online advocacy

This paper fits within the Identity and Online Advocacy because it explores how identity can be attacked and abused within an online sphere and through social media.

Introduction

Identity when put in an online atmosphere can be precarious. In some ways it is beneficial, but this paper explores what happens when identity is attacked and what the social media platforms fail at doing to prevent real life consequences. As many people within the Muslim community have faced. Hate crime is not something new, especially when it manifests as hate speech. Defined by the United Nations as “any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor” in other words, hate speech is when someone is attacked and broken down based on their identity or an aspect of their identity (2019, p. 2). These attacks on minorities have become more prevalent with online media; Zachery Laub explains that social media allows for individuals that are “inclined toward racism, misogyny, or homophobia” to have their views reinforced through these platforms. Further, a correlation between online media posts and violence against minorities has been investigated and research into what causes this violence as well (2019). Within a sample of fifteen to thirty year old’s in the United Kingdom, United States, Finland and Germany it was found that on average forty-three per cent had encountered hateful material online (Williams et al, 2019, p. 95). This is especially prevalent with online Islamophobia, which is a very common. This exposure to online hate speech, especially at this scale, is bound to cause some forms of apathy towards seeing it, and generally can be ignored. This influx of online hate speech is only monitored to a certain extent, although attempts are made by the platforms—who claim the behaviour is unacceptable—struggles occur with the systems they attempt to use. Overall, social media has led to hate crime and hate speech—especially towards Muslim people—being something integrated into everyone’s lives, but due to the lack of monitoring and repercussions online from the social media platforms, such as Facebook and Twitter, has led to real life consequences and actions;

nevertheless, the online policing is lacking and causing people to go unchecked, a prime example explored will be the Christchurch Mosque Massacre.

Facebook

Identity on Facebook is intricately connected to the person, requiring the use of a full name and birthday along with other facts about yourself such as: your religion, family members, relationship status, interests and workplace. Facebook's online presence is huge, having approximately a third of the world's population active on Facebook, but although they attempt to control what can or is said online, their methods are flawed and have been shown to lead to real life hate crime (Laub, 2019). Facebook's own statement on hate speech states that their belief is when people online don't feel attacked for their identity, they are more likely to share their voices and "that's why we don't allow hate speech on Facebook. It creates an environment of intimidation and exclusion, and in some cases may promote offline violence" (n.d.). This acknowledgement of online hate speech also implies that they have rules and regulations against the perpetration and continuous publication of hate speech, this implication means that those who are a part of hateful groups or participate in hate speech are facing consequences of their actions. But this isn't always true. Richard Allan, the VP EMEA Public Policy, explains that Facebook is committed to "removing hate speech any time" they become aware, explaining that the previous two months 66 000 posts were reported per week (2017). Although Allan admitted to their system not being perfect because of ambiguity, context, intent and human mistakes (2017). Despite their regulations and attempts to keep hate speech from being posted many Islamophobic posts are created and, in some cases, sparks real violence. With new social media developments, the ability to find niche areas that create "echo chambers", which perpetrates the same ideas, therefore reinforcing them, becomes so much stronger and harder to regulate (Muller & Schwarz, 2020, p. 2). The Online Hate Prevention Centre reported that they "searched over fifty different Facebook pages" and found 349 instances of hate speech "directed towards Muslims" (Awan, 2014, p.145).

For example, in Germany, antirefugee sentiment has grown exponentially since the "Alternative for Germany (AfD)" had created their Facebook page and consistently posted about their views on refugees (Muller & Schwarz, 2020, p. 3). In particular, a correlation occurred with how many hate mongering posts about refugees occurred and attacks on refugees such as arson, physical attacks, property damage and vandalism (Muller & Schwarz,

2020, p. 17). Further, outages across Germany which disrupt the exposure to the platform, is shown to lower the amount of hate crimes within particular areas by approximately ten per cent (Gupta, 2020). Jahnvi Gupta further explains, on Muller and Schwarz's research, that "municipalities with Alternative fur Deutschland users were three times more likely to experience an attack" during the period of AfD anti-refugee hate speech posts (2020). Alongside this correlation of attacks connected to posts, when Facebook outages occurred the probability of a hate crime lowered by twelve per cent (Gupta, 2020). This shows, not only the relation between online hate crime affecting the real world, but how little monitoring Facebook does of the page. Further, despite their desire to protect identities and cultures of people using their platform, political identities in this situation are being prioritised over the individuals. With the ability for such a significant page to continue their aggressive posts about Muslim people, Facebook should prioritise the people being attacked. Real life hate crimes occur when online hate speech, within niche pages on Facebook, are more prevalent; although, Facebook is not the only prevalent social platform that has neglected their duties in preventing hate speech.

Twitter

Twitter's online presence is also significant, but their lack of control over hate speech online has also progressed to real life violence. Twitter, at the end of 2020, had 353 million users and the market hitting an all-time high in 2020 of \$40 billion compared to the \$23.04 billion approximately a year beforehand (Iqbal, 2021). Overall, Twitter's online presence is significant, because of this their policing of online hate speech is lacking due to the mass influx of posts every day. Their personal rules are: "Hateful conduct: You may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories" and the consequences for violating their policy, depending on the individual's past record and the severity, can be from having to only use Twitter on a read only mode for a portion of time to account suspension (Twitter, 2020). These consequences seem to still allow the continued use of Twitter, therefore not providing any intervention between the

user and others within their niche echo chambers. This does nothing to prevent the pursual of real-life violence and further attacks on Muslim people.

Just like Facebook, hate speech is prevalent on Twitter. For example, after the Woolwich attack in south-east London where a British soldier was murdered by two men who were Muslim, the increase in Islamophobic attacks surged and although many of these were offline a great amount of people vented their frustrations online (Awan, 2014, p. 133). They were ultimately blaming all Muslims and grouping them all within the same stereotype of violent offenders (Awan, 2014, p.133). At the time, Imran Awan, expressed the frustration in the lack of research on Islamophobia online, this is despite the fact that many people, including Muslim government officials, have expressed their concern and discomfort at the emerging Islamophobic sentiments (2014, p. 137). The issue went further by remaining unaddressed by Twitter, who expressed that the victim of the abuse should “block someone or close their account” (Awan, 2014, p. 138). This reliance on the user is dangerous and ineffective in preventing further action, it doesn’t stop the attacker from continuing their tirade it only prevents the victim from seeing it. Twitter also made a statement saying they cannot stop what isn’t reported, thus creating the issue of reliance on the individuals online to report the content in spite of their previous statement telling the victim to block the attacker (Awan, 2014, p. 138). This can cause many issues, specifically in social platforms niche communities and echo chambers which allow like-minded people to group together; therefore, nothing will be reported in these cases. Although there are attempts made to control the hate speech online, there is only so much that can be done, although how these sights should be policed is a focus of many people’s interests.

Development of policing online hate

Within the communities being attack, especially the Muslim community, a demand for online policing of these issues has arisen. This has developed excessively since March 15th 2019, when armed gunmen “opened fire into two mosques in Christchurch, New Zealand” (Evans, 2019). The gunmen went by Brenton Tarrant on Twitter and had posted photos of the weapons two days before the massacre, further they livestreamed the rampage to Facebook (Evans, 2019). These gunmen were also active on a number of groups on Facebook that expressed Islamophobic sentiments (SBS, 2021). The gunmen were avid internet users, particularly Facebook; Twitter, and their favourite, 8Chan, describing before the attack that

they were going to take their ‘shit-posting’ to ‘taking the dialogue offline (Williams et al., 2019, p. 97). Further from the leadup of hateful commentary that predicted the attack, the seventeen-minute video was taken down within the hour, but was reuploaded over two million times on ‘Facebook, YouTube, Instagram and Twitter’ (Williams et al., 2019, p. 97). These sites flooded with praise for the shooters and opened a discussion from the Muslim community to these social media platforms, in particular The Australian Muslim Advocacy Network (AWAN) has been campaigning for Facebook to act against ‘the spread of hate speech and dangerous conspiracy theories directed towards Muslims, on various accounts, links and groups on the site’ (SBS, 2021). These social media websites reliance on ‘artificial intelligence, user reporting and staff known as content moderators’ has led to inconsistencies of application depending on the language, clarity of the rules and accessibility (Laub, 2019). Issues occur with these methods through faulty algorithms, niche communities not reporting content, moderators who are burdened with the volume and severity of the posts they sift through and lack of resources (Laub, 2019).

Beyond the success or failure of Facebook’s monitoring, anti-Muslim hatred online is still prominent. Including pages made for Muslim people receiving regular hateful comments or posts, beyond that anti-Muslim pages the issues with this combination of hate crimes and the social media posts going viral is this hatred becomes the standard and normalised within communities (Oboler, 2016, p. 45). This new version of online hate has made it significantly more difficult for online social media platforms to track and shut down (Oboler, 2016, p. 46). One difficulty identified is that there are challenges that arise from identifying what is hate speech, whether it is simply a criticism of religion versus what is hateful and vicious (Oboler, 2016, p. 58). The way for improvement can come in a variety of ways, particularly in improving the regulations on not just the platforms level, but further up including government and police. This adjustment can potentially help ease the burdens on moderators, rely less on participants and hopefully improve the clarity of rules for the artificial intelligence.

Conclusion

Ultimately, the issues with Islamophobia online has grown to an insurmountable extreme. It is important for us to understand not just the consequence of the online hate speech but further to understand how these online platforms are failing their communities. Identity is something cherished by most who use the internet, the ability to be whoever you want online is a beautiful concept. Except those who are extremist, who take their dream of eradicating an

entire community online to speak with others who agree. The Muslim community online has faced innumerable challenges online with their identity not being linked to themselves but instead their entire community. It is important to be able to recognise what is wrong with these systems and platforms, so then there are the options to speak out further and encourage more action from the platforms that facilitate these people.

References

- Allan, R. (2017, June 27). *Hard questions: Who should decide what is hate speech in an online global community?* About Facebook. <https://about.fb.com/news/2017/06/hard-questions-hate-speech/>
- Awan, I. (2014). Islamophobia and Twitter: A typology of online hate against Muslims on social media. *Policy & Internet*, 6(2), 133-150. <https://doi.org/10.1002/1944-2866.poi364>
- Awan, I., & Oboler, A. (2016). The Normalisation of Islamophobia through Social Media: Facebook. In *Islamophobia in cyberspace: Hate crimes go viral* (1st ed., pp. 41-61). Routledge.
- Evans, R. (2019, March 15). *Shitposting, inspirational terrorism, and the Christchurch mosque massacre*. *bellingcat*. <https://www.bellingcat.com/news/rest-of-world/2019/03/15/shitposting-inspirational-terrorism-and-the-christchurch-mosque-massacre/>
- Facebook. (n.d.). *Community standards*. https://www.facebook.com/communitystandards/hate_speech
- Gupta, J. (2020, October 7). *Study suggests connection between social media, hate crime*. *Hindustan Times*. <https://www.hindustantimes.com/more-lifestyle/study-suggests-connection-between-social-media-hate-crime/story-bOrJ4O7ZtcGkPgvVddUabO.html>
- Iqbal, M. (2020, June 23). *Twitter revenue and usage statistics (2020)*. *Business of Apps*. <https://www.businessofapps.com/data/twitter-statistics/>
- Laub, Z. (2019, April 11). *Hate speech on social media: Global comparisons*. *Council on Foreign Relations*. <https://www.cfr.org/backgrounder/hate-speech-social-media-global-comparisons>

- Müller, K., & Schwarz, C. (2020). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*. <https://doi.org/10.1093/jeea/jvaa045>
- SBS News. (2021, March 21). *Australian group takes fight to Facebook, saying platform is 'awash' with hateful Islamophobia*. <https://www.sbs.com.au/news/australian-group-takes-fight-to-facebook-saying-platform-is-awash-with-hateful-islamophobia>
- Twitter. (2020, September 22). *Hateful conduct policy*. Help Center. <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy>
- United Nations. (2019). *United Nations Strategy and Plan of Action on Hate Speech*. <https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20Plan%20of%20Action%20on%20Hate%20Speech%2018%20June%20SYNO%20PSIS.pdf>
- Williams, M. L., Burnap, P., Javed, A., Liu, H., & Ozalp, S. (2019). Hate in the machine: Anti-Black and anti-Muslim social media posts as predictors of offline racially and religiously aggravated crime. *The British Journal of Criminology*. <https://doi.org/10.1093/bjc/azz049>