

Exploring The Negative Impacts of Online Hate Speech on The Identity of Black Community on Twitter.

Abstract

The purpose of this paper is to briefly discuss about the negative impacts of hate speech that is being spread on Twitter towards the Black community. This paper will outline the different stages towards the construction of these hatred rhetoric and further discuss on how online hate has constructed the term Black Twitter and a certain kind of hashtags known as Black Tagging. The different negative effects on the self- esteem and social identities of the Black community that these online hate are propagating will also be explored.

Keywords: #HateSpeech, #BlackTwitter #Blacktags #Anonymity #BlackCommunity #SocialIdentities #OnlineIdentities

The rise of social media platforms has revolutionized how people interact with each other. The decentralization aspect of these platforms has made it ideal for individuals to be able to produce as well as exchange experiences, knowledge, issues and their different perspectives. However, in defiance of their democratic attributes, the involvement in reinforcing emerging kind of hatred online that they have, is receiving more scrutiny lately (Davidson, Warmley, Macy & Weber, 2017). Since individuals are more interactive than it has ever been, they have established a distinct visual identity on the platforms that represents them and their foundation. This is considered to be beneficial as they establish a connection with their peers, families as well as information. However, when this aspect is threatened, a specific facet of their lives is endangered. The black community is getting targeted with hate speech online, mainly on Twitter, which is impacting their sense of self and is in turn contributing to individuals refusing the disclosure of that personality (Nakamura & Chow-White, 2012). As a result, this is significantly impacting their self-esteem, complicated by the fact since they are being confronted by hatred speech that they probably lack adequate handling capabilities. This essay will delve into the negative perceived effects that hate speech in creating on the self-esteem and social identity of black community using mainly Twitter as an approach. The essay will explore different aspects on the sources of hatred online and how it contributes to the creation of online hate speech with relation to the rise of 'Black Twitter' and the development of certain 'hashtags' and its negative impacts.

Twitter and Online Hate Speech Towards Black Community

As Twitter is continuing to resemble social and economic environments in different regions around the globe, emerging racial trends are becoming more common. These racial rhetoric emerge with deceptive techniques like using memes and dummy profiles so as to provoke hate and racism. Researches show concrete data that due to the phenomenon of anonymity, the sense of being anonymous has a significant impact on users' behaviour. Individuals are relatively more violent in these cases where they believe they are unidentified. Therefore, these haters communities are more at ease sharing their

hatred since they won't be found out or face the repercussions. According to Burnap & Williams (2015), Twitter has been considered the networking site having the highest amount of hate speech towards the Black community due to the huge proportions of anonymous accounts. However, despite certain users having the lack of anonymity, there have been notable examples of hate speech through their real profiles and details on Twitter (Chaudhry, 2015). For example, Waseem and Hovy (2016) investigated a total of 130k tweets on Twitter out of which 16.8k were concluded to be hateful towards the Black community. With the expansion of this platform, structured racist groups are formed and Twitter is used as a medium to spread hate speech (Johns 2017). Apart from this, racial websites are created and users are redirected from Twitter to these websites attracting an increasing amount of participants, which therefore expands and reinforces a racist group identity and provoke abuse towards the Black community (Parenti 2013). Since digital sites are ideal for the personalization of texts according to targeted audience, therefore Twitter allows these racial communities to be sufficiently characterized in the virtual world, through delivering effective methods to connect a larger number of participants and propagate vitriolic dialect. This promotes a feeling of belonging within themselves.

Researchers are putting much effort into identifying and recognizing the virtual hateful rhetoric, however, relatively limited is known concerning its proliferation on conventional or radical sites and the frequency of hatred changes when it comes to the field's issues. Beauchamp (2018) identified that creators of sexist and abusive expression appear to approach in a smoother, rather implicit manner then progressing towards a pernicious manner. According to the authors, some of this is attributed to declining degree of interpersonal exclusion as these people move into a more severe virtual community. Through evaluating members who emit violent rhetoric to others who do not, Beauchamp (2018) identified abusive participants as being strongly interconnected to each other. In consequence, the abusive information produced by these individuals are spread quicker, further and to a broader demographic in comparison with participants not engaged in spreading hatred. Therefore, due to these movements, this has become the main reason as to why hate speech is being redirected more visibly on Twitter.

Black Tagging & Black Twitter Through Online Hate Speech on Twitter

Individuals are shifting to justifying media outlets as democrats become more fragmented. Therefore, this creates an escalating feeling of ethical detachment that contribute to denigration, although this tendency does not exactly reflect sufficiently for constructive interaction due to the varied perspectives and vocal expressions needed for a stable democratic state (Tillery, 2019). Some examples that can be considered are when Barack Obama was re-elected and when athlete Jason Collins made public that he was homosexual (Kian, 2015). Such occurrences provoked hate speech on Twitter about racial and gender identity. Within this context, the projection of hate speech on Twitter is perceived as convoluted, societal issue, with different morphologies in various geopolitical contexts. With regards to hate speech composition, structures of toxic effects and violence are determined by how specific individuals are attacked (Wilson, 2011). As a result, hatred towards a community can evolve into various types of violence progressively or merge with antagonism for other identities. Participants tend to identify

groups sharing the same views when they meet others whose ideologies challenge their self, setting in motion a structure of discrimination, whereby groups not formerly thought to be clearly distinguishable are now categorized as the "other" (Boyd, 2013). Therefore, as additional "others" are discovered, the challenge of identification grows, necessitating a restraint towards them.

According to studies, hate speech proliferation is predicted by a systemic characteristics of Twitter that relate to the capabilities and restrictions of conceptual data (Mahendran, 2007). These viral political hashtags on Twitter depend upon various interactions that are considered to trigger a wider group of individuals across many segments of the platform. Therefore, to effectively comprehend this situation in this context, reverting towards recognition and corresponding those hashtags towards a specific community, i.e the Black community's ethnic origins is problematic (Sharma, 2013). Through envisioning visible discrepancies in online communities regarding political hashtags, research on the trend of racial hashtag propagation continue (Wilson, 2011). As a result, a presumption concerning socially constructed racist consumer attitudes persists. These results aren't strictly applicable to all sorts of hashtags. Still, they do call into question some superficial notion of Blacktags as a personification of Black people by them to this particular dividing behavior through Black-tagging and Black Twitter (Sharma, 2013). Northrup (1989) explains it as a phobia prolonged yet pointless presence, but destructive "to the self as a physical threat." For example, if a central sense of self is questioned by another individual, a challenged is established.. A confrontational posture as evidence attacks their self-esteem and identity which as a consequence, leads to an advanced level of deterioration (Bromley 2011). Since this demonstrates the importance of social interactions and the proliferation of racist groupings, this discloses for what cause Blacktags are so widely spread; due to the fact that hatred tweets conveying political criticisms in a vitriolic form of commentary that is aligned with the Black community are curated and disseminated as viral contents (Foxman & Wolf, 2013). Hence, hate speech on Twitter has lead to a more de-territorialized ethnic community of Black users who fear becoming a divided community dominated by a range of geopolitical initiatives. The online interpretation of Twitter produces the negative effects and interpretations of Blacktags.

Impact on Self-Esteem and Social Identities of Black Community

Although verbal abuse is among the causes contributing to communal violence, it has a significant impact on the intensification of emotions of collective hatred. According to Watanabe, Bouazizi & Ohtsuki (2018), this is especially accurate in the virtual environment, by which the privacy of internet interaction will lead to participants expressing increasingly harmful views than they would normally. Conflicts between groups are compounded as participants begin to feel that standard social norms should not adhere to them. In the same manner, hatred expression on these social sites creates a substantial barrier among the participants and the leaders, empowering respondents to self-interpret without fear of retaliation (Watanabe, Bouazizi & Ohtsuki, 2018). Izsak (2015) describes social identity as an individual's sense of self to a community, as well as the sentimental significance that comes with it. Individuals have

several intertwining identities connected with a precisely established and important social classes along with vague definitions like ethnic origin. According to the social identity theory, connections to socially constructed identities are a critical indicator of how individuals identify and view themselves (Tillery, 2019). Therefore, when an individual associates with a community, their self-esteem is influenced by how they are viewed by the community. Hateful rhetoric on social media can trigger anxiety, especially among traditionally oppressed or vulnerable groups. According to other findings, such disclosure can cause individuals to remove themselves from civic conversation, be it on the net or offline, hurting freedom of expression and community involvement (Chan, Ghose & Seamans 2015). Evidently, qualitative data suggest that vulnerability of racism has almost identical effects as being the victim of hateful talks, such as mental distress and community terror (Watanabe, Bouazizi & Ohtsuki, 2018). Humanitarian activists have concluded that failing to track and fight hateful rhetoric digitally will intensify marginalized victims' subjugation, rendering them more susceptible to exploitation, thus rendering the bigotry communities more responsive to this kind of hostility (Burnap & Williams, 2015).

Demaneing words or actions may have serious consequences for participants of these directed communities. Although this association is not inherently definitive, the already targeted individuals to the harmful rhetoric tend to have a lower commitment to their families and show greater concentrations of dissatisfaction (Chetty & Alathur 2018). Internet susceptibility to offensive rhetoric is often linked to a long-term lack of constitutional discussion. Harmful rhetoric on the internet has engendered inter-group divisions in a number of ways, contributing to offline incidents and fracturing social stability in some cases (Izsak 2015). For example in the US, the tragic events of Breonna Taylor and George Floyd demonstrated the mistreatment that the Black community encounters. These atrocities revealed the concerns that Black community endure due to online online discrimination that contribute to offline homicides (Cappelli, 2020). Hatred will weaken the meaningful facets of the self that are extracted from group identity, robbing participants of their integrity, by violating the effective attributes of belonging to a social circle. According to studies, becoming a target of the hatred speech on the internet will trigger the target's fundamental identification of becoming associated with insecurities; rather than associating oneself with admirable qualities, the victimized individual of hatred will associate with their group with threatened feelings (Müller & Schwarz, 2020). This makes them doubt their own values as representatives of virtual communities and societies.

Nevertheless, certain evidence suggests that representatives of communities whose personal identity are often threatened through hatred and shame are less prone to developing concern to the haters while others actually dismiss it (Watanabe, Bouazizi & Ohtsuki, 2018). Gerstenfeld (2017) discovered that certain persons who are despised are better at attributing the hatred than those who are rarely threatened. Any participants of victimized communities, for example, are said to embrace marginalization throughout time. The sense of self arises from "reflected appraisals," in which individuals of marginalized communities are mindful of their victimization and integrate derogatory views within their consciousness (Root, 2018). Furthermore, even though individuals of

vilified communities are ignorant of their victimization, they may behave in compliance with subconscious biases depending on how they are viewed by outside discrimination, rather than their own flaws, nullifying the impact of hatred.

Conclusion

To conclude, this paper has explored the way hate speech on Twitter has negatively impacted the self and social identities of the Black community. This paper has briefly discussed about the origin of the online hatred, how online hate on Twitter has given rise to terms like Black Twitter and Black Tagging and the different stages that help construct these hatred rhetoric. Further research includes the exploration of how hate speech towards a specific community can be detected online and automatically removed as a way to safeguard the Black community against racial proliferation on social networking sites

Reference List

- Beauchamp, N., Panaitiu, I., & Piston, S. (2018). Trajectories of hate: Mapping individual racism and misogyny on Twitter. Unpublished working paper.
- Boyd, D. (2013). White flight in networked publics. How race and class shaped American teen engagement with MySpace and Facebook. In L. Nakamura & PA Chow-White (Eds.), *Race after the internet*, 203-222.
- Bromley, R. (2011). Beast, vermin, insect—Hate media and the construction of the enemy: The case of Rwanda, 1990—1994. In Billias, N., Praeg, L. (Eds.), *Creating destruction: Constructing images of violence and genocide* (pp. 40–59). Brill/Rodopi. <https://doi.org/10.1163/9789401200387>
- Burnap, P., & Williams, M. L. (2015). Cyber hate speech on twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & internet*, 7(2), 223-242
- Cappelli, M. L. (2020). Black Lives Matter: The Emotional and Racial Dynamics of the George Floyd Protest Graffiti. *Advances in Applied Sociology*, 9(10), 323.
- Chan, J., Ghose, A., & Seamans, R. (2015). The Internet and racial hate crime: Offline spillovers from online access. *MIS Quarterly*, 14(2), 381–403
- Chaudhry, I. (2015). # Hashtagging hate: Using Twitter to track racism online. *First Monday*.
- Chetty, N., & Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression and Violent Behavior*, 40, 108–118

- Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. In Proceedings of the International AAAI Conference on Web and Social Media, 11(1).
- Foxman, A. H., & Wolf, C. (2013). *Viral hate: Containing its spread on the Internet*. Macmillan.
- Gerstenfeld, P. B. (2017). *Hate Crimes: Causes, Controls, and Controversies*. London: Sage.
- Izsak, R. (2015). Hate speech and incitement to hatred against minorities in the media. UN Humans Rights Council. arXiv.org. www.ohchr.org/EN/Issues/Minorities/SRMinorities/Pages/Annual.aspx
- Johns, A. (2017). "Flagging White Nationalism 'After Cronulla': From the Beach to the Net." *Journal of Intercultural Studies* 38 (3): 349–64.
- Kian, E. M. (2015). A case study on message-board and media framing of gay male athletes on a politically liberal web site. *International Journal of Sport Communication*, 8(4), 500-518
- Mahendran, D. (2007). The facticity of blackness: A non-conceptual approach to the study of race and racism in Fanon's and Merleau-Ponty's phenomenology. *Human Architecture: Journal of the Sociology of Self-Knowledge*, 5(3), 18.
- Müller, K., & Schwarz, C. (2020). From hashtag to hate crime: Twitter and anti-minority sentiment. Available at SSRN 3149103.
- Nakamura, L., & Chow-White, P. A. (2012). Introduction—race and digital technology: Code, the color line, and the information society. *Race after the Internet*, 1, 7-24.
- Northrup, T. A. (1989). The dynamic of identity in personal and social conflict. In Kriesberg, L., Northrup, T. A., Thorson, S. J. (Eds.), *Intractable conflicts and their transformation* (pp. 55–82). Syracuse University Press.
- Parenti, M. (2013). Extreme right organizations and online politics: A comparative analysis of five Western democracies. In P. Nixon, R. Rawal, & D. Mercea (Eds.), *Politics and the Internet in Comparative Context* (pp. 155–173). London: Routledge.
- Root, E. (2018). Staging scenes of co-cultural communication: Acting out aspects of marginalized and dominant identities. *Communication Teacher*, 32(1), 13-18.
- Sharma, S. (2013). Black Twitter? Racial hashtags, networks and contagion. *New formations*, 78(78), 46-64.

- Tillery, A. B. (2019). Tweeting racial representation: how the congressional Black Caucus used Twitter in the 113th congress. *Politics, Groups, and Identities*.
- Waseem, Z., & Hovy, D. (2016, June). Hateful symbols or hateful people? predictive features for hate speech detection on twitter. In *Proceedings of the NAACL student research workshop* (pp. 88-93).
- Watanabe, H., Bouazizi, M., & Ohtsuki, T. (2018). Hate speech on twitter: A pragmatic approach to collect hateful and offensive expressions and perform hate speech detection. *IEEE access*, 6, 13825-13835.
- Wilson, J. (2011). Playing with politics: Political fans and Twitter faking in post-broadcast democracy. *Convergence*, 17(4), 445-461.