

The Impact of Online Hate Speech on Muslim Women: some evidence from the UK Context

Kyriaki Topidi

in: Matthias C. Kettemann (ed.), How Platforms Respond to Human Rights Conflicts Online. Best Practices in Weighing Rights and Obligations in Hybrid Online Orders (Hamburg: Verlag Hans-Bredow-Institut, 2022)

The Impact of Online Hate Speech on Muslim Women: some evidence from the UK Context

Kyriaki Topidi

EUROPEAN CENTRE FOR MINORITY ISSUES

“Technology is neither good or bad; nor is it neutral.”¹

Introduction

The recent ECRI General Policy Recommendation No.5 on preventing and combating anti-Muslim racism and discrimination (ECRI, 2022) in its preamble draws a troublesome picture on the features of anti-Muslim discrimination patterns, including in their digital dimensions. It notes ‘the prevalence of false accusations affecting Muslim communities as a whole without distinguishing between Muslims and religiously disguised extremists’ (ECRI, 2022, 7). It also points out the intersectional dimension of anti-Muslim hatred and the ‘growing gendered abuse against Muslim women, particularly those that publicly manifest their faith’ (Ibid, 7).

Starting, therefore, from continuing attempts in Europe to draw an ‘essentialized’ picture of Muslims as a group characterized by foreignness, backwardness, threat and cultural incompatibility with core European values (Ibid, 10), this contribution focuses on the implications of a ‘racialized’ representation of Muslims online, placing emphasis on women. An intersectional analysis of the impact of online hate speech is justified by the fact that this category of persons experience harm as a consequence of the interaction of their gender, religious affiliation and ethnic origin as identity markers that single them out as ‘targets’ for online hate speech.

Against the backdrop of the exponential growth of online hate speech targeting Muslims,² the aim of the analysis is to illustrate, using evidence from the UK context where data is to some extent available, the dimensions of a gendered reading of anti-Muslim hatred online, especially in connection to the harm that such speech creates. By securitizing Muslim communities, online hate speech sets that basis for serious limitations of the freedom of expression, freedom of association and political participation of a vulnerable segment of contemporary European societies (ECRI, 2022, 21).

Digital Hate and its Impact: Intersectionality as a Methodological Lens

The concept of *intersectionality* is linked to the assumption that an individual has multiple identities. As a conceptual frame, it suggests that by adopting an intersectional approach, research abandons single-axis

¹ Melvin Kranzberg, “Technology and History: ‘Kranzberg’s Laws’”, *Technology and Culture*, 27, No.3 (1986): 545.

² See indicatively the results of the 2021 Council of Europe survey of the Secretary General’s Special Representative on Antisemitic, anti-Muslim and other forms of religious intolerance and hate crimes.

perspectives (e.g. gender, ethnicity, religion, social class, etc.) in order to explain more comprehensively an individual's experiences and socio-legal positionality (Crenshaw, 1991). The consideration of multiple demographic categories has been initially devised to describe ethnic minority women and their employment prospects through the notion of disadvantage and oppression. More than that, these categories have served as evidence of intersecting systems of power (in the form of racism and sexism) that create social inequalities (Collins, 2015). Previous research has addressed intersectionality as a paradigm (Bilge, 2010; Hancock, 2007), as a theory of generalized identity (Nash, 2008) as well as a methodological approach (McCall, 2005; Yuval-Davis, 2006).

Intersectionality, however, as proposed by Crenshaw, has the potential to provide a frame for two additional sets of inquiries: first, how the individual effects of discrimination on the basis of single characteristics separately (e.g. gender and ethnic origin) have cumulative effects through their intersection on right-holders; and second, the implications of applying intersectionality against the risk of over-generalisation (Hancock, 2007). More recently, on the basis of Holvino's (2010) framework of *simultaneity*, intersectional studies have proposed to take account of the multiple levels of norms, processes and structures that reproduce and maintain inequality, often in historically informed patterns of systematic disadvantage. In this light, the intersectionality of gender, ethnicity and socio-economic position are shown to create variations among women, even when they belong to the same ethnic group (Acker, 2006; Holvino, 2010). When analyzing the impact of online hate speech on Muslim women in the UK, the intersectional analytical trajectory matters because it highlights two neglected aspects: first, the intersectional process in itself (i.e. the interplay of several multiplicative factors of vulnerability such as gender, religion or race) and second, the extent to which there is conflation of intragroup differences.³

In general terms, the variations in which discrimination is manifested depend on the combination of various identity markers and are neither easily captured by statistical data nor are they addressed within anti-discrimination legislation. In stricter legal terms, no explicit legal references to intersectional discrimination exist within the European Convention on Human Rights, with a limited exception under Article 14 ECHR which has the potential to address intersectional discrimination.⁴

At the individual level, an intersectional lens acknowledges the violation of one's individual right to equal treatment but at the structural level, if ignored, it reinforces discrimination within and between legally protected categories. This implies that structural forms of discrimination operate on the basis of deep-seated social hierarchies and call for more contextual analysis. So far, the individual dimension of discrimination has prevailed in the European context but is inadequate to address the socio-legal impact on online hate speech on Muslim women (CIJ/ENAR, 2019, 13). This is because such an approach makes abstraction of the broader context of discrimination (i.e. the societal system that is favourable to hate speech online), as well as of the role played by other relevant actors/institutions in mapping the effects of laws and policies on the question.

The aim of the study is therefore to stress the need to highlight intragroup differences among Muslim women online against over-generalisations. Reversing or mitigating the harm of online hateful expression

³ Both dimensions are directly sourced from K. Crenshaw's work (1989, 1993).

⁴ See also the B.S. v Spain case finding a violation of Article 3 and 14 ECHR for an isolated application of intersectionality. EU law and the CJEU are also silent on the issue. See a contrario the UN CEDAW General Recommendation No. 28, para.18 that stipulates: "The discrimination of women based on sex and gender is inextricably linked with other factors that affect women, such as race, ethnicity, religion or belief, health, status, age, class, caste and sexual orientation and gender identity. Discrimination on the basis of sex or gender may affect women belonging to such groups to a different degree or in different ways to men."

caused to these women, as an exercise in inclusivity, is nevertheless complex. McCall (2005, 1782) has argued in favour of an intra-categorical approach when applying an intersectional lens because it allows to 'uncover the differences and complexities of experience' at the intersection of multiple categories. At the same time, the study will also extend the intersectional analysis beyond the 'target' subjects (i.e. Muslim women) to address the role of other powerful actors with the ability to influence the degree/type of harm caused online (McBride et al., 2015, 13). Ultimately, the justification of an intersectional approach responds to the need to engage with the 'missing' voices, moving beyond the acknowledgement of their absence in the debates. More than that, the goal is to question the relationship between race, gender and religion in the online space in order to address the processes of marginalization and discrimination at play.

In terms of identity practice, intersectionality scholarship has already established in the UK context, that regardless of their individual characteristics, visible gender and religious/ethnic markers activate stereotypes against immigrant women, in particular Muslims, emphasizing their perceived submissiveness and weakness (Hwag & Beauregard (2021), 8; Ali et al. (2017); Kamenou (2008); Wyatt and Silvester (2015)), as well as their invisibility.

The internet and female Muslim identity in the UK: A brief background note

The increase of online Islamophobic incidents and hate speech in Britain is related to the rise of far-Right discourse, as well as to Brexit and lastly to the role of media news and social media in normalizing hateful acts (Hopkins et al, 2020, 6).⁵ As such, a clearer link is gradually being established in this respect between the general economic, social and political climate and the online ecosystem. The majority of female Muslims in recent surveys (80 per cent) live in wide-ranging fear of experiencing Islamophobia (physically or verbally) (Hopkins et al., 2020, 12-13), often with repercussions on their family members (e.g. the victims' children at school).

Such fear based on gendered Islamophobia is well documented in the UK context: for example, Arab or South Asian origin increases the perception of the victims as 'risks' in responses in NE England (35.3 and 25.8 per cent respectively), while wearing a headscarf or having brown skin also attracts fear among respondents (94.6 and 80.2 per cent respectively) (Hopkins et al., 2020, 14). Experiences of fear, exclusion and immobility are therefore far from uncommon among Muslim women. In that sense, online hate speech targeting Muslim women operates against an offline environment which is already characterized by tense community relations, rooted within more entrenched forms of racism and discrimination in society. The consolidation of social media in the hands of a limited number of actors has further affected the ways in which information and expressions are shared and received by their users, with implications for this particular category of users (Aguilera-Carnero and Azeez, 2016, 22).

The UK is one of the few European countries that collect data about ethnic minorities. Protection against racism in the UK extends to the online domain and therefore incitement to racial and religious hatred or violence is also illegal online (Feldman & Littler, 2014). Despite some prosecutions under the Public Order Act (1986), targeted discrimination and prejudice online are often related to 'trigger events' around which online hate speech incidents increase drastically (PRISM, UK Report, 9, stating the example of the assassination of British soldier Lee Rigby in 2013 by two Muslim men.) A second important feature relates

⁵The Report indicates that a majority of the Muslim respondents surveyed in North-East England (84.3 per cent) had experienced directly Islamophobia, with 15.5 per cent having experienced online abuse (Ibid, at 7).

to the active involvement of far-right parties in spreading hostile messages online, with data suggesting an important majority of hate speech online incidents linked to the British far right (Feldman et al., 2013: 21). The use of social media by far-right parties is essentially designed to facilitate the flow of their opinions and arguments, also through mainstream media and reach their supporters (PRISM UK Report, 11).

Those with 'visible' Muslim identity online are targeted, often repeatedly, and by the effect of such repetitive targeting, the line between the offline and online impact of hate speech becomes blurred, especially in connection to intimidation and abuse (Awan & Zempi, 2015,6). Hostile comments, racist posts, memes and images, often circulated through fake ID profiles, constitute the basic framework for hateful online expressions, particularly around the 'trigger events'.⁶ The use of pictures magnifies the range of abuse and suffering of victims online (Awan & Zempi, 2015, 20). Within this context, it is worth highlighting how race and religion become interlinked in attacking Muslim women online in the UK (Awan & Zempi, 2015, 13) and end up reinforcing each other. Stigmatization, 'othering' and stereotypization happens through abusive and provocative language often including offline threats against the victims and their families. However, moving beyond the victim/perpetrator frame in the case of online hate speech, may have long term benefits in combatting online hate speech.

In the UK, religious hate crime figures declined by 5 per cent in 2020, and still, Islamophobic hate constituted 50 per cent of total incidents (Home Office, 2020). At the same time, it is becoming clearer that Islamophobic online discourses impact hate crimes, within and beyond the UK context (Levin, 2016; Mueller and Schwarz, 2020). In a sense, it can be argued that both online and offline hate speech represent the evidence of the internalization of certain cultural discourses propagated by some politicians too (Burnett, 2017). Far right and populist groups' online discourses function as enabling factors towards the creation of digital ecosystems reinforcing exclusion (Perry and Scrivens, 2018). They do so by promoting narratives of danger and untrustworthiness (Home Office, 2020).

In 2016, the Home Secretary established the National Online Hate Crime Hub and in 2017 a Home Affairs Select Committee ran an inquiry on hate crime leading to a Law Commission review seeking to address the inadequacy in legislation on online hate speech. More recently, OFCOM (2018) found in the UK an expansive development of hate speech online, with almost half of UK internet users seeing hateful content in the past year, although around 50 per cent of anti-Muslim hate was only produced by 6 per cent of users, often classified as politically anti-Islam (Demos, 2017).⁷ The question on the correlation between online hate speech on race/religion with police recorded religiously aggravated crimes remains positive (Williams et al., 2020) in Britain, although it is unclear whether online hate speech precedes or follows offline hate crime.⁸

The online (technological) context also appears to perpetuate 'climates of unsafety' in the vicious cycle of victimization (Williams et al., 2020, 113). In this sense, algorithmic tools towards content moderation and classification of hate speech can only be finite in their contribution towards the reduction of hate speech, especially as cultural and linguistic context of online discourse shifts.

⁶ See for example the #KillAllMuslims twitter hashtag after the Paris shooting in January 2015.

⁷ According to Hope Not Hate (2019) figures, 5 out of 10 far rights social media activists with the biggest online reach in the world were British.

⁸ Other factors such as demographic factors, unemployment rates, etc matter as well (Williams et al., 2020, 112).

The English Legal Framework on Online Hate Speech

Hate speech, for the purposes of the present study, is understood as “covering all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, including intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants and people of immigrant origin.” (Council of Europe, 1997).

In the digital space, however, the concept of ‘cyber hate’ or its co-terminous ‘hate speech online’ are of particular focus in the present context and refer to “any use of electronic communications technology to spread anti-Semitic, racist, bigoted, extremist or terrorist messages or information. These electronic communication technologies include the internet (i.e. websites, social networking sites, Web 2.0 user-generated content, dating sites, blogs, online games, instant messages and e-mail) as well as other computer – and cell phone-based information technologies (such as text messages and mobile phones).”⁹

It is an offence, in England and Wales, to incite hatred through hate content based on the grounds of race, religion and sexual orientation. The dissemination of the hateful material can happen through words, pictures, video music and can take a variety of forms such as messages calling for racial/religious violence, glorification of violence on the basis of race/religion or within chat forums where people are invited to commit hate crimes.

While there is no specific legal definition available on online hate speech in the British context, the Crown Prosecution Service has provided some guidance on the qualification which will determine prosecution: speech which is motivated by any form of discrimination/hostility against the victim’s ethnic or national origin, gender, disability, age, religion or belief, sexual orientation or gender identity can form the object of criminal prosecution.¹⁰ In fact, according to Home Office Data, between 2017 and 2018, there was a 40 per cent increase in online offences flagged as hate crimes.¹¹

While no single piece of hate crime legislation exists, a number of offences, based on hostility aimed at one of the ‘protected characteristics’ (e.g. race, religion, sexual orientation)¹² provide for harsher sentences. Legislative references in this respect include the Crime and Disorder Act 1998 and the Criminal Justice Act 2003 which consider hostility as an aggravating factor of an offence.¹³ A distinction can be made between hatred and hostility with the former being more serious (Williams, 2019, 35) and carrying a public order dimension through attacks targeting an entire group. No legal definition of hostility is available. Other related offences to hate speech include communications’ offences and harassment (e.g. messages that are offensive/indecent, menacing or false).

⁹ Anti-Defamation League, Responding to Cyberhate – Toolkit for Action, http://www.adl.org/internet/Bindel_final.pdf

¹⁰ On a European level, the approach of the CPS echoes to some degree the general definition of hate speech within the Council of Europe’s 1997 Recommendation (97) 20 which defines hate speech : “(...) as covering all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, including: intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants and people of immigrant origin.”

¹¹ The number of offences rose to 1605 in 2018, although the data available were treated with caution [Cf. Home Office, Hate Crime England and Wales 2017/18 Report].

¹² According to CPS, protected characteristics are identified within “[a]ny criminal offence which is perceived by the victim or any other person, to be motivated by hostility or prejudice, based on a person’s disability or perceived disability; race or perceived race; or religion or perceived religion; or sexual orientation or perceived sexual orientation or a person who is transgender or perceived to be transgender” (Williams, 52).

¹³ Hostility in both Acts is framed as follows: at the time of committing the offence or immediately before or after doing so, the offender demonstrated towards the victim hostility based on protected characteristics (see e.g. section 145 and 146 of the CJA 2003).

Additionally, within the Public Order Act 1986 are punished offences that stir up racial and religious hatred. Stirring religious hatred occurs when someone says something, including online, which is threatening, with the intention to stir up religious hatred. Stirring hatred, in this specific legislative context, should target an entire group, and not simply one person, and pose a threat to public order. Put simply, this offence entails more than voicing an opinion or causing offence.¹⁴ Finally, the Communications Act 2003 and the Malicious Communications Act 1988 cover the sending of offensive, menacing, indecent or obscene messages and treat them as offences with no requirement for the communication to be received so that sending it suffices. Section 1 of the MCA 1988 criminalizes the sending of communications to another person with the purpose of causing 'distress and anxiety'. The message must also be indecent, grossly offensive, a threat or false information. Section 127 of the CA 2003 similarly carries the requirement of the message to cause 'inconvenience or needless anxiety to another' in order to qualify under the offence. The CA 2003 excludes communication sent over a 'private' network. For both Acts, the terms of the offences are generally ambiguous (e.g. the notion of 'gross offensiveness').¹⁵ Overall, in practice, the majority of online hate speech expressions are pursued under one of the communications offences.

In criminal law terms, it is the targeted nature of (online) speech that stigmatizes its victims, along with an eventual incitement of others to hate in a threatening/abusive way, that distinguishes it from offensive or controversial speech, otherwise protected under Article 10 ECHR (Williams, 2019, 14-15). It is worth comparing this approach with the definitions of online hate speech adopted by social media platforms. For example, the 2019 Twitter Rules and Policies approach hate speech as online speech that "(...) promote(s) violence against or directly attack[s] or threaten[s] other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability or serious disease". According to the same Twitter rules "accounts whose primary purpose is inciting harm towards others on the basis of these categories" are also not allowed. Twitter's Policy also forbids hateful imagery (e.g. symbols associated with hate groups or images depicting others as less than human or non-human). Similarly, the 2019 Facebook Community Standards approach online hate speech on the basis of three tiers of severity, according to their degree of offensive character and impact of the speech in question.¹⁶

In terms of conceptualization, the negative impact of online hate speech on its victims is identified in both the English criminal law context as well as within the tech companies' terms and policies of use. The effectiveness in identifying and removing as well as prosecuting such expressions, because of their harmful impact is less established. Tech companies are still failing to remove large quantities of such speech¹⁷, while the criminal law system's response is limited as the reporting of online hate offences is low.¹⁸ The emerging positive correlation between hate speech with religiously aggravated offences on the streets confirms the salience of the issue (Muller and Schwarz, 2017, 2018; Williams et al., 2019).

¹⁴ Section 29J POA 1986 stipulates that nothing in the Act "(...) prohibits or restricts discussion, criticism or expressions of antipathy, dislike, ridicule, insult or abuse of particular religions, or the beliefs or practices shall not be taken of itself to be threatening."

¹⁵ The Law Commission, *Abusive and Offensive Online Communications: A Scoping Report*, London: Crown Copyright, 2018 at 6.

¹⁶ Tier 1 characterizes violent, dehumanizing speech; Tier 2 implies inferiority, contempt/disgust or uses slurs; Tier 3 speech excludes or negatively targets people (reference?).

¹⁷ See the UK Digital, Culture, Media and Sport Committee's Report on 'Disinformation and "fake news"' (2019) on the limited results in preventing such damaging expressions.

¹⁸ See the 2018 Inspectorate of Constabulary and Fire and Rescue Service Report 'Understanding the Difference: The Initial Police Response to Hate Crime' that stressed the inadequacy on online offence responses.

The Role of Platforms in spreading hate against Muslim Women

Law is not, however, an exclusive or even determinant tool for regulating online hate speech: other important parameters of regulation are market considerations, social norms and the architecture of the internet per se. The actors with the power to control what is possible and available online indirectly contribute in shaping how users behave in a more visible way. Social media -and media in general – have difficulty to depict Muslim women as neither oppressed nor dangerous (ENAR, 2016,3). Presented as not having agency, such women appear online most often as victims, usually involved in incidents or legal procedures (ENAR, 2016, 13 citing evidence from Italy, France, Denmark).

More specifically, the cycle of hatred is perpetuated by the role of social media within which are also reflected the negative representations of Muslims on broadcast television. Social media have also been blamed for allowing the normalization of xenophobia and propaganda disinformation about minority groups more generally (Evolvi, 2018; Farkas et al. 2017). The use of Twitter in this context as a platform in relation to anti-racist politics is indicative¹⁹: more recently, popular social media platforms have allowed the circulation of gifs and memes perpetuating stereotypes (Sharma, 2013), racism and misogyny (Ringrose,2018) likely to escape moderation and ultimately portraying minority groups as shared enemies (Wahl- Jorgensen, 2019, 110). In a sense, such use of the platforms suggests that these latter are likely to enhance the agency of groups of haters that are well-organized and with stronger ties among them (Poole et al., 2021, 1438). In addition, two dominant streams fuel this cycle further: the online discourse of a number of politicians and the political system and the contribution of far-right political views in normalizing such discourse.

When it comes to content moderation, remaining neutral is not an option for platforms: attempts to do so have led to increasing amounts of disinformation, threatening democracies and inter-group relations (Suzor, 2019, 22). The initial refusal of technology providers to deal with hate speech in its socio-legal dimension belongs, therefore, to a by-gone era. Abuse and harm contained in content hosted by the platforms in question are tightly linked to the architecture of commercial internet and cannot be exclusively placed in the remit of state criminal law systems. This happens also because, often, online abuse may not raise to the standard required to qualify as a criminal law offence and yet produces harm that leads to the exclusion of entire groups of people from opportunities offered by the internet (Suzor, 2019, 31).

Online abuse is conducive to the silencing of minority voices. While major social media networks prohibit online hate speech, as mentioned, the effectiveness of such internal rules can be questioned. From the platforms' perspective, internal rules against hate speech are either not enforced or are applied in discriminatory ways, depending on the identity/position of the speaker. Even when platforms enforce such rules, people spreading hate against minorities can choose to act in groups, instrumentalizing platforms' reporting functions to silence further minority groups (Suzor, 2019, 36-37), with women being a widespread target (Ibid, 146). As fundamentally, online hate speech in the form of name-calling, threats or bullying constitute both a cause and effect of a broader web of inequality (Suzor, 2019, 32). This is because such speech contributes and/or perpetuates the disempowerment and lack of autonomy of those that it targets.

Well-known techniques for screening content online include preemptive filtering, post-publication review and automated detection. All three of them carry the implications of either filtering insufficiently or

¹⁹ Twitter's use is also linked to the decline of investigative journalism within mainstream media and its emergence as a source of news in itself (Poole et al., 2020, at 1417).

exceedingly online materials that impact the capacity of speech of minority groups. Without a full understanding of how platforms are moderating content, it becomes challenging to identify and reverse patterns of exclusion. The use of flagging moderation systems combined with blocking and/or filtering tools are therefore not sufficient to reverse the worrying trend. Relying on users to flag breaches of internal rules on hate speech is also not enough as in a good proportion of cases, the content flagged mirrors dominant biases in society (Suzor, 2019, 143). Ultimately, online abuse and hate speech become much more than an issue of content of classification: they are essentially the reflection of a culture of systemic discrimination that platforms should not encourage.

In connection to the last point, in the UK context, a YouGov poll found that 74 per cent of the British public knew little or nothing about Islam, with 64 per cent getting their information from the media (ENAR, 2016, *Forgotten Women- UK National Report*). Similarly, UK Gallup found that 30 per cent of the British public perceives the hijab as a threat and 16 per cent would not want a Muslim as a neighbour (ENAR, 2016, *UK National Report*). The visibility of an individual's Muslim identity, both offline and online, is key to anti-Muslim hate crime and yet the victims of online anti-Muslim hate are less inclined to report incidents to the criminal justice system (Awan & Zempi, 2020, 3).

At the same time, Facebook, Twitter and Instagram include considerable amounts of anti-Muslim hate speech against women (Awan, 2014). In light of the main features of anti-Muslim hate speech online targeting women, ICTs can no longer be treated in isolation, from both a normative and a regulatory perspective, but rather as embedded in a society (Schroder & Ling, 2013, 790). After all, the market and public pressure influence the policies of platforms (Suzor, 2019, 108), beyond strict legal considerations.

The consideration of the broader context leading to the production of online hateful speech may provide an alternative starting point to understand and then contain hate speech online. Instead of tackling individual expressions online, it could be useful to focus also on the context of online speech production, with emphasis on what an individual user is producing, sharing or engaging with (Vidgen et al., 2021, 16). A behavioural approach to online hate can be relevant when designing the responses to hate speech (e.g. quarantines, warning, counter speech, etc), especially given the limited number of 'online haters' producing the bulk of anti-Muslim hate speech. A more articulate understanding of the impact of such speech on its victims can also guide further in making decisions on resource allocation, especially in connection with victim support (Vidgen et al., 2021, 16). It should be noted nevertheless that patterns of online Islamophobic expressions vary across time, in quantity and quality, although it remains relatively consistent that a limited number of users is responsible for most hateful content online (Vidgen et al., 16).

Typology and impact of anti-Muslim hate speech: Gendered Experiences

Hate speech in general terms undermines the pursuit of social cohesion and inclusion within multicultural societies. Using vilification messages against group identities, it transforms attacks against ideas to attacks against persons. Common themes within Islamophobic content encountered online include attacks on moderate Islam on the basis of distrust towards Muslim communities in general; arguments that Muslims are not a race, but a religion and therefore vilification against them cannot be racist; or calls for the wiping out of Muslim culture (Oboler, 2013, 18-20). A basic typology of anti-Muslim hate speech, according to Oboler's categorization (2013, 14 et seq.), treats Muslims as security threats (e.g. Muslims as terrorists), as cultural threats (e.g. availability of halal food), as economic threats (e.g. Muslims as 'drains' to welfare

systems), as violent threats as well as objects of dehumanization (e.g. Muslims compared to animals) and demonization (e.g. collective allegations of criminality).

Muslim women's experiences online are equally highly contextual and varied although in direct link and continuity with the offline world. Consequently, a conceptualization of online communications that begins and ends online would be largely outmoded (Wellman & Hampton, 1999). More than that, according to Varisco (2010), "there is no pure Islamic presence, separated from other relevant forms of identity, in cyberspace any more than there is (...) in the real world (Ibid, 176).

Still, within social media, conceived as networks of interaction, anti-Muslim online hate speech acquires differentiated characteristics when compared with offline forms: anonymity, immediacy, impact and reach amplify the impact of hateful expressions (Brown, 2018). The following features, according to UNESCO (2015), differentiate online harmful expressions from offline ones: their permanence, as speech can move across platforms and be repeatedly linked; their itinerancy, across various online spaces; the possibility to be anonymous when expressing hate and finally the transnational reach of online hate speech. This differentiated impact applies especially to 'casual' forms of Islamophobia that make space for more divisive and extreme forms and expression of hate (Vidgen et al., 2021, 3). Far-right actors are particularly involved in creating 'walls of hate' (Awan, 2016),²⁰ although available research on online hate speech suggests that it varies across time, users, context and geography (Ganesh, 2018). This may be due to online spaces being decentralized, complex and disrupted (Margetts, John, Hale, Yasseri, 2015).

The visibility of the veil online amplifies its symbolism offline as a reminder of women's 'powerlessness, vulnerability and oppression' (Chakraborti and Zempi, 2012, 280). At the same time, the veil operates as an identification tool towards Islamophobic victimization. The issue of Muslim female 'visibility' is in itself complex: on one hand perpetrators rely on the symbolic visibility of such women to identify them as targets (e.g. visual signs such as the veil) but at the same time, online representation of the same group is victimized in largely invisible ways being represented as without agency, escaping regulation and public condemnation (Allen, 2015, 299). In sum, Muslim women's online visibility as victims of hate speech only serves the goal of 'erasure' in social value and relevance (Sayyid, 2003).

From notions of safety, physical and emotional security as well as self-worth, the impact of collective victimization has been 'escaping' the radars of legal and policy responses (Bowling, 1999; Chakraborti, 2010). In fact, Victim Support Data (2013) demonstrated that the degree of impact of crime on its victim is not necessarily commensurate to the gravity of the crime in criminal law terms. Ultimately, hate crimes, including online hate speech, are not directed towards an individual victim only but also against the community to which they are presumed to belong (Chana, 2020, 73). Due to this dual dimension, online hate speech can be framed as being beyond a reaction to 'trigger' events. It can also be considered as forming part of a wider cycle of systemic victimization, stigmatization and intimidation (Bowling, 1999, 18).

The 1997 Runnymede Trust Report on *Islamophobia: A Challenge for Us All* had identified in an earlier context the impact of Islamophobic incidents as detrimental on personal and community identities, including on one's individual and community worth (Allen, 2015, 289). Later in 2005, the Open Society Institute's Report *Muslims in the UK: Policies for Engaged Citizens* (Choudhary, 2005) highlighted the

²⁰ For example, among Islamophobic tweets sent between 2016/17, 15 per cent were sent by 1 per cent of the most active users and 50 per cent by the 6 per cent of most active users (Demos, 2017). See also Awan and Zempi (2017) showing how far right actors exploit virtual environment and world events to incite hatred.

emergence of stereotypes about Muslims and Islam, attributed in the public sphere to those most visible. Building on the observation of how dominant representations of Muslim women as oppressed/subjugated are, the Report showed the parallel development of a sense of justification developed among perpetrators attacking veiled Muslim women as representative of all Muslims, ignoring their individual dimensions and position as humans and subjects of legal rights/citizens.

'Islamophobia', a neologism encountered in 1996 within the Runnymede Trust established Commission on British Muslims, is premised on a monolithic view of Islam based on an essentialized interpretation of the communities as inferior and alien to Western values (Aguilera-Carnero et al., 2016, 24). It has been defined as 'indiscriminate negative attitudes or emotions directed at Islam or Muslims' (Bleich 2011). Cyber-Islamophobia, of particular interest here, is present through both blogs and social media but equally within online traditional media outlets. Cyber-hate, as such, covers abusive online material, which can and does lead in some cases to offline violence, cyber violence, cyber stalking as well as online harassment, through the use of images, videos and text. Anti-Muslim hate is premised on hostility based on a person's real or perceived Muslim religion (Awan, 2014) and is presented in the English case, as English patriotic speech which in practice dehumanizes Muslims (Awan & Zempi, 2015).

Normatively, online hate speech embodies the digital clash between freedom of expression, autonomy, human dignity and ultimately equality (Gargliardone, 2014). In other forms, it can also be ambiguously presented, without hateful language but disguising hate in its tone and/or context. The emotional and behavioural impact of online hate speech can be both short term (e.g. isolation, resentment, shame, anger) and/or long term (e.g. low self-esteem, development of prejudice against hate speaker's group, concealment of identity) (Williams, 17).

Against women in particular, the internet is creating opportunities for new ways, tools and means to perpetrate crimes (Banks, 2001, 163). The most common forms of online abuse against women include their depiction as unintelligent, hysteric or ugly (Jane, 2014), which are increasingly normalized due to lack of removal, prosecution²¹ and reporting. Earlier empirical data from Twitter, for example, indicate the presence of stereotyping against Muslims (Aguilera-Carnero et al., 2016) as violently opposed to other religions as well as secularism. The use of linguistically generic references to demonize the entire community is also very present, creating new patterns of racism online (Van Dijk, 2000).

The harm that derives from online hate speech is well captured by Waldron's definition of the term as 'profound disrespect, hatred, and vilification of the members of minority groups' (Waldron, 2012). In fact, the normative contradiction between hate speech and free speech can be filtered through the focus on the harm created by certain categories of speech: Benesch in her work has introduced the concept of 'dangerous speech' precisely to express the amplifying violence by one group against another (Benesch, 2013,1). Labelled as a 'gendered crisis', affecting disproportionately women (Elmir, 2016), it is beyond the focus of this analysis, however, to debate the advantages/disadvantages of the criminalization of hate speech in the British legal order. Precisely because of its impact and given the multi-stakeholder constellation relevant for online content moderation, the online space operates as an unequal environment of suppression, with visible patterns of hierarchy of human worth. It operates against the background of 'online misogyny', a term employed to signify a form of gendered types of 'trolling' (also called e-bile) that is facilitated and

21 Online abuse against women tends to be tackled only when escalating to death threats or when abuse continues 'offline'; (Jane, 2014).

normalized through digital platforms. (Ringrose, 2018).²² According to the Law Commission Scoring Report on Abusive and Offensive Online Communication, the qualitative effect of hate speech has a wide range and includes psychological effects, emotional harms and physiological harms, exclusion from the online space, economic harms and wider societal harms affecting groups and social inclusion more broadly.²³

Vulnerability is thus an integral part of Muslim women's lives due to the growing normalization effects of hateful expressions (and crimes) online. This leads them to operate under the weight of continuous risk assessments both online and offline in order to minimize, and if possible, avoid further harassment (ENAR, 2016, 26). Online abuse also results in attempts by victims to become less 'visible' by removing their veils (Zempi & Chakraborti, 2014; Perry & Alvi, 2012) so as to decrease their vulnerability. Perceived as attacks on their identity and with implications for their self-esteem, belonging and safety, the impact of anti-Muslim hate crime may exceed that of other types of crime (Awan & Zempi, 2015, 25). Verbal abuse and hate speech remain the most common incidents quantitatively.²⁴ Hate images and posts are, within this context, characterized by loaded generalisations (ENAR Report, 2016, 27). The effects of these gendered offences extend to self-censor online (with implications for free speech exercise); vulnerability; anxiety or terror (Lewis- Hastelen, 2011). This type of hate speech therefore suggests the connection between such vitriolic and harmful expression with the silencing and exclusionary patterns operating against female online (and offline) agency, self-control and ultimately empowerment (Barlow & Awan, 2016, 7). Put simply, "perpetrators (usually men) attack Muslim women not only because they belong to an ethnic, racial or religious minority, or because they are women, but because of their combined intersectional identities" (ENAR, 2016, 31).

The primary implications of virality of such content lead to self-censorship as mentioned,²⁵ and as crucially to the creation and perpetuation of digital infrastructures for the promotion of hate (Malik, 2017). These structures rely on both online and offline expression avoidance strategies by victims, although ironically, anti-Muslim expressions online are framed often as freedom of expression instances (Ekman, 2015). Visual representations of 'Muslimness' through clothing serve to signal potential targets for anti-Muslim hatred, within a flattening process of racialization by proxy of religious affiliation (Ganesh, 2016, 32).

Such harm caused by hate speech, according to Waldron (2012), is often left out of balancing exercises in regulatory and policy assessments of responses to such expressions. And yet, "members of the vulnerable groups targeted are expected to live their lives, conduct their business, raise their children (...) in a social atmosphere poised by their sort of speech (...)."

Moving Forward: The Law's Response and examples of Best Practices

Internet governance proposals place a disproportional focus on the legal aspects connected to states' contributions in legal rules on content moderation (Suzor, 2019, 113). Criminal law systems, constituting an integral part of the regulation of online hate speech, are less adequately designed, however, in the UK context (and elsewhere) to address online abuse and even less the disproportionate targeting of specific

²² The example of Twitter is illustrative as a structure which allows public users to address each other but also do keyword and hashtag searches.

²³ The Law Commission, *Abusive and Offensive Online Communications: A Scoping Report*, London: Crown Copyright, 2018, at 4.

²⁴ Tell MAMA 2013-14 data found that 82 per cent of online incidents were about verbal abuse/hate speech.

²⁵ Hampton et al (2014) have found that self-censorship affects by priority those holding minority opinions.

groups such as women and minority religious groups.²⁶ The UN Special Rapporteur on violence against women noted in 2018 the “inadequate and substandard responses from intermediaries concerning online gender-based violence.”²⁷ Moving forward, adopting human rights standards for platforms, as corporate entities, is a way to reduce risks (and negative press) while improving current exclusionary practices. In simpler terms, for platforms, taking on human rights responsibilities signifies taking responsibility for the consequences of their choices on users (Suzor, 2019, 131). This implies the introduction of monitoring systems assessing the effects of their policies and practices on the human rights of their users. Beyond privacy concerns and free speech, the impact of hate speech is fundamentally a choice on the type of equality these companies endorse. Disregarding social context and impact of hate speech shows preference for a formal type of equality, over a substantive one.

In the UK context, state supported attempts to address anti-Muslim online hate speech can be conceived on two levels: as institutional and as normative.

Institutional Responses

According to the UK Law Commission’s initial 2018 report related to Offensive Online Communications,²⁸ online hate crime has an acknowledged strong misogynistic dimension with damaging effects which the current criminal law content does not capture (Williams, 2019, 44). The following institutional initiatives are some of the measures introduced recently to mitigate the impact of online hate speech, including on women, with varying degrees of state involvement:

- The National Online Hate Crime Hub, established in 2017, functions institutionally as a central point for the collection of all reports of online hate crime. It is designed to improve the victim experience and increase rates of successful prosecution (Williams, 2019, 43), especially given the prevailing reluctance of victims of online hate to report incidents. Alongside the Online Hate Crime Hub, the UK Council for Internet Safety (UKCIS) caters inter alia for women and girls victims of online harms and hate speech.²⁹
- Third-party reporting centres have been put forward as an instrument to overcome mistrust towards the police felt by members of religious and racial minorities and a barrier to reporting further hate crime incidents (Macpherson, 1999). Such centres were introduced to improve accessibility of the criminal justice system for all (Home Office, 2016). In practice, their establishment, granted only to existing organizations, created an asymmetric web of reporting venues with mitigated impact (Chana, 2020, 75-76).
- ‘Tell MAMA’ (Measuring Anti-Muslim Attacks) is a service launched in February 2012. It has been facilitated by the civil society organization Faith Matters. Tell MAMA records instances of Islamophobic hate crimes and incidents, including those happening online. It has been the first entity of its kind in Europe and is supported by the government in Britain. Victims can report

²⁶ The Law Commission, *Abusive and Offensive Online Communications: A Scoping Report*, London: Crown Copyright, 2018, at 8-9.

²⁷ Dubravka Simonovic, ‘Report of the Special Rapporteur on Violence against Women, its Causes and Consequences on Online Violence Against Women and Girls from a Human Rights Perspective’, UN General Assembly, June 14, 2018, at para. 73.

²⁸ The Law Commission, *Abusive and Offensive Online Communications: A Scoping Report*, London: Crown Copyright, 2018.

²⁹ <https://www.gov.uk/government/organisations/uk-council-for-internet-safety> .

incidents through freephone numbers, by email, SMS, Twitter, Facebook and receive support by trained staff. Tell MAMA passes information to the police using an online reporting system.³⁰

- In one of its earlier reports, covering the period between April 2012-April 2013, TellMAMA noted that 74 per cent of Islamophobic incidents registered (out of a total of 584) happened online, targeting Muslim women in considerable proportion (58 per cent of the total number of which 80 per cent concerned veiled/visibly recognizable Muslim women) (Copsey et al, 2013). The report also noted low levels of reporting of incidents to the police (with 63 per cent of all incidents going unreported). In a more recent 2017 Report, TellMAMA found the continuation of the increase on online Anti-Muslim incidents.

Normative Responses

The 2019 White Paper on Online Harm, published by the Department for Digital, Culture, Media and Sports and the Home Office proposed a new statutory duty of care for tech companies. The proposed legislation (Online Harm Bill) is intended to ‘bring much needed clarity to the regulatory landscape (...)’³¹ through proportionate, risk-based and tightly defined means of implementation, moving beyond liability models. This is necessary given the great diversity of online services and harms it proposes to cover.³² To meet the ‘duty of care’ the companies concerned³³ will need to understand the risk of harm to individuals on their services and put in place systems to improve their safety.³⁴ It also requires companies to set codes of practice, establish transparency as well as effective /accessible complaint mechanisms for users and the obligation to collaborate with law enforcement in cases of illegal online harms. The current Bill also categorizes content as illegal, harmful and legal but harmful. The notions of harm and harmful content are not, however, clearly defined, although references to harm provoked through existing criminal offences (e.g. hate crime) are covered. This is likely to be insufficient as ‘harmful’ content is highly context dependent. The fulfilment of the tech companies’ duty is to be overseen by an independent regulator, already designated as being Ofcom. The regulator in question would dispose of a range of powers against companies breaching the duty of care that include the issuing of fines,³⁵ the disruption of their activities and the liability of individual members of the senior management of such companies.

The draft legislation is partially responsive to online abuse and intimidation suffered by women and those from minority backgrounds,³⁶ as it embraces equality concerns, in particular misogyny and persecution of minorities. It has been criticized nevertheless for its implications on freedom of speech as well as for

³⁰ <https://tellmamauk.org>.

³¹ UK Government, Online Harms White Paper: Full government response to the consultation, December 2020, Command Paper Number 354, Crown Copyright, at para. 17.

³² Range of services covered include social media services, consumer cloud storage sites, video sharing platforms, online forums, dating services, online instant messaging services, peer-to-peer services, videogames which enable interaction with other users and online marketplaces.

³³ Only companies with direct control over the content and activity on a service will owe the ‘duty of care’.

³⁴ Means of redress offered to companies are foreseen to be content removal, sanctions against offending users, reversal of wrongful content removal or sanction or changes to company processes and policies.

³⁵ Fines of up to 18 million pounds or 10 per cent of a company’s annual turnover (whichever is higher) can be imposed.

³⁶ UK Government, Online Harms White Paper: Full government response to the consultation, December 2020, Command Paper Number 354, Crown Copyright, see Box 9 -Anonymous Abuse.

uncertainty on the treatment of content that is not illegal but still harmful. Evidence-based considerations of harm caused by digital content also remain underdeveloped.

References

- Aguilera- Carnero, C. and Azeez, A.-H. (2016), 'Islamonausea, not Islamophobia': The many faces of cyber hate speech, *Journal of Arab and Muslim Media Research*, 9(1), 21-40.
- Ali, F.; Malik, A.; Pereira, V. et al. (2017) A relational understanding of work-life balance of Muslim migrant women in the West: Future research agenda, *International Journal of Human Resource Management*, 28(8): 1163-1181.
- Allen, C. (2015) 'People hate you because of the way you dress': Understanding the invisible experiences of veiled British Muslim Women victims of Islamophobia, *Int'l Review of Victimology*, Vol.21 (3), 287-301.
- Allen, C.; Isakjee, A.; Young, O. (2013) "Maybe we are hated": The experience and impact of anti-Muslim hate on British Muslim women, Institute of Applied Social Studies.
- Awan, I. (2014) Islamophobia on Twitter: A typology of online hate against Muslims on social media, *Policy and Internet*, 6, 133-150.
- Awan, I. and Zempi, I. (2015), Virtual and physical world anti-Muslim hate crime, *The British Journal of Criminology*.
- Banks, K. (2001), Leave the internet alone, gender equality and ICT, *APWIN Journal*, 3, 147-173.
- Barlow, C. and Awan, I. (2016) "You need to Be Sorted Out with a Knife": The Attempted Online Silencing of Women and People of Muslim Faith within Academia, *Social Media and Society*, October-December 2016: 1-11.
- Benesch, S. (2013) Dangerous Speech: A proposal to prevent group violence: The dangerous speech project, <http://www.worldpolicy.org/content/dangerous-speech-along-the-path-to-mass-violence>
- Bowling, B. (1999), *Violent Racism: Victimization, Policing and Social Context*, Oxford: Oxford University Press.
- Burnett, J. (2017) Racial violence and the Brexit state, *Race and Class*, 58: 85-97.
- Center for Intersectional Justice /ENAR, (2019) Intersectional discrimination in Europe: Relevance, Challenges and Ways Forward, https://www.intersectionaljustice.org/img/intersectionality-report-FINAL_yizq4j.pdf
- Central for Intersectional Justice, Intersectionality at a Glance in Europe – Fact Sheet, April 2020, available at https://www.intersectionaljustice.org/img/2020.4.14_cij-factsheet-intersectionality-at-a-glance-in-europe_du2r4w.pdf
- Chakraborti, N. and Zempi, I. (2012) The Veil under attack: Gendered dimensions of Islamophobic Victimization, *Int'l Review of Victimology*, 18(3): 269-284.
- Chana, S. (2020) 'Working Towards a Better Understanding of Islamophobia', *British Journal of Community Justice*, Vol. 16(2), 72-91.
- Chetty, N.; Altathur, S. (2018) Hate Speech review in the context of online social networks, *Aggression and Violent Behavior*, 40: 108-118.
- Choudhury, T. (2005) *Muslims in the UK: Policies for Engaged Citizens*, London: Open Society Institute.
- Copsey, N.; Dack, J.; Littler, M. et al. (2013), *Anti-Muslim Hate Crime and the Far Right*, Middlesbrough: Centre for Fascist, Anti-Fascist and Post-Fascist Studies.
- Crenshaw, K. (1989) Demarginalizing the Intersection of Race and Sex, *University of Chicago Legal Forum*, 139-167.
- Crenshaw, K. (1993) Mapping the margins: Intersectionality, identity politics and violence against women of color, *Stanford Law Review*, 43: 1241-1299.
- DEMOS (2017), *Anti-Islamic Content on Twitter*
- Ekman, M. (2015) Online Islamophobia and the politics of fear: Manufacturing the green scare, *Ethnic and Racial Studies*, 38(11): 1986-2002.
- Elmir, R. (2016) Muslim women bear the burden of Islamophobia, *Gulf News: The Views*, 18 September.

- European Network Against Racism (ENAR) (2016), *Forgotten Women: The impact of Islamophobia on Muslim women*, <https://www.enar-eu.org/Forgotten-Women-the-impact-of-Islamophobia-on-Muslim-women/>
- Evolvi, G. (2018) #Islamexit: Inter-group antagonism on Twitter, *Information, Communication and Society*, 22(3): 386-401.
- Farkas, J.; Schou, J. and Neumayer, C. (2017) Cloaked Facebook pages: explaining fake Islamist propaganda in social media, *New Media and Society*, 20(5):1850-1867.
- Feldman, M. and Littler, M. (2014) Tell MAMA Reporting 2013/14 – Anti-Muslim Overview, Analysis and ‘Cumulative Extremism’
- Gagliardone, I; Gal, D.; Aleves, T.; Martinez, G. (2015) *Countering Online Hate Speech*
- Ganesh, B. (2016) *The Geography of Anti-Muslim Hatred, Annual Report 2015*, London: Tell MAMA.
- Hampton, K.; Shin, I. and Lu, W. (2016) Social media and political discussion: when online presence silences offline conversation, information, communication and society, 1468
- Home Office (2016), *Action Against Hate: The UK Government’s Plan for Tackling Hate Crime*, London: Home Office.
- Home Office (2020) *Hate Crime, England and Wales: 2019 to 2020*, <https://www.gov.uk/government/statistics/hate-crime-england-and-wales-2019-to-2020>
- Hope Not Hate (2019), *State of Hate 2019, Hope Not Hate*
- Hopkins, P; Clayton, T. and Tell MAMA (2020), *Islamophobia and Anti-Muslim Hatred in North- East England*, <https://www.tellmamauk.org/wp-content/uploads/2020/06/ISLAMOPHOBIA-AND-ANTI-MUSLIM-HATRED-IN-NORTH-EAST-ENGLAND-090620.pdf>
- Hwang, S. and Beauregard, A.T. (2021), Contextualising intersectionality: A qualitative study of East Asian female migrant workers in the UK, *Human Relations*, 1-26.
- Jane, E. (2014) “You’re an ugly, whorish, slut”: Understanding e-bile, *Feminist Media Studies*, 14, 531-546.
- Jubany, O. and Roiha, M., (2016) *Backgrounds, Experiences and Responses to Online Hate Speech: A Comparative Cross-Country Analysis*, PRISM <https://sosracismo.eu/wp-content/uploads/2016/07/Backgrounds-Experiences-and-Responses-to-Online-Hate-Speech.pdf>
- Kamenou, N.; Netto, G. and Fearfull, A. (2013) Ethnic minority women in the Scottish labour market: Employers’ perceptions, *British Journal of Management* 24(3): 398-413.
- Kim, J. (2019) Ethnic capital, migration and citizenship: A Bourdieusian perspective, *Ethnic and Racial Studies*, 42(3): 357-385.
- Law Commission, *Abusive and Offensive Online Communications: A Scoping Report*, London: Crown Copyright, 2018.
- Levin, B. (2016) *Special Status Report: Hate Crime in the United States*, San Bernardino: California State University.
- Lewis- Hasteley, H. (2011) “You should have your tongue ripped out”: The reality of sexist abuse online, *New Statesman*, 3 November 2011.
- Macpherson, W. (1999) *The Stephen Lawrence Inquiry: Report of an Inquiry by Sir William Macpherson of Cluny*, HMSO, London.
- Malik, N. (2017) London won’t let hate prevail, *Gulf News: The Views*, 25 March
- Mc Bride, A; Hebson, G.; Holgate, V.J. (2015) Intersectionality: Are we taking enough notice in the field of work and employment relations?, *Work, Employment and Society*, 29(2), 331-341.
- McCall, L. (2005) The complexity of intersectionality, *Signs – A Journal of Women in Culture and Society*, 30(3): 1771-1800.
- Muller, K. and Schwarz, C. (2017), ‘Fanning the Flames of Hate: Social Media and Hate Crime’, Working Paper, University of Warwick.
- Muller, K. and Schwarz, C. (2018) ,Making America Hate Again? Twitter and Hate Crime Under Trump’, Working Paper, University of Warwick.
- Muller, K. and Schwarz,C. (2020), From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment, SSRN, 1-47.
- Murphy, A. (2021) Political Rhetoric and Hate Speech in the case of Shamima Begum, *Religions*, 12: 834
- Oboler, A. (2013) *Islamophobia on the Internet: The growth of online hate targeting Muslims*, Online Hate Prevention Institute Report, IR 13-7,

- <https://www.researchgate.net/publication/271706783> Islamophobia on the Internet The growth of online hate targeting Muslims
- OFCOM (2018), News Consumption in the UK: 2018, OFCOM.
- Pennington, R. (2018) Making Space in Social Media: #MuslimWomensDay in Twitter, *Journal of Communication Inquiry*, Vol.42(3), 199-217.
- Perry, B. and Alvi, S. (2012) "We are all Vulnerable": The Terrorism Effects of Hate Crime, *Int'l Review of Victimology*, 18(1): 57-71.
- Perry, B. and Scrivens, R. (2018) A Climate for Hate? An Exploration of the Right-Wing Extremist Landscape in Canada, *Critical Criminology*, 26: 169-187.
- Poole, E.; Giraud, E.H.; de Quincey, E., (2021) Tactical interventions in online hate speech: The case of #stopIslam, *New Media and Society*, Vol.23(6), 1415-1442.
- Ringrose, J. (2018) Digital feminist pedagogy and post-truth misogyny, *Teaching in Higher Education* 23(5): 647-656.
- Runnymede Trust (1997) Islamophobia: A Challenge for Us All, London: Runnymede Trust
- Sayyid, B.S. (2003) A Fundamental Fear: Eurocentrism and the Emergence of Islamism, London: Zed Books.
- Schroder, R. and Ling, R. (2013) Durkheim and Weber on the Social Implications of New Information and Communication Technologies, *New Media and Society*, 16(5), 789-805
- Seglow, J. (2016) Hate speech, dignity and self-respect, *Ethical Theory and Moral Practice*, 19(5): 1103-16.
- Sharma, S. (2013) Black Twitter? Racial hashtags, networks and contagion, *New Formations*, 78:46-64.
- Suzor, N. (2019) Lawless, Cambridge University Press, 2019
- UK Government (2020) Online Harms White Paper: Full government response to the consultation, December 2020, Command Paper Number 354, Crown Copyright
- UNESCO (2015) Countering online hate speech
- Van Dijk, T. (2000), New(s) Racism: A Discourse Analytical Approach, in Cottle, S. (ed.) Ethnic Minorities and the Media, Buckingham: Oxford University Press, 33-49.
- Varisco, D.M. (2010) Muslims and the Media in the blogosphere, *Contemporary Islam*, 4, 157-177.
- Victim Support (2013), How Crime can affect you, <http://www.victimsupport.org.uk/help-for-victims/how-crime-can-affect-you>
- Vidgen, B.; Yasseri, T.; Margetts, H. (2021), Islamophobes are not all the same! Study of far rights actors across Twitter, *Journal of Policing, Intelligence and Counter Terrorism*
- Wahl-Jorgensen, K. (2019) *Emotions, Media and Politics*, London: Polity Press.
- Waldron, J. (2012) The Harm in Hate Speech, London: Harvard University Press.
- Williams, M. (2019) Hatred Behind the Screens: A Report on the Rise of Online Hate Speech, Mischcon de Reya, <https://hatelab.net/wp-content/uploads/2019/11/Hatred-Behind-the-Screens.pdf>
- Williams, M.L.; Burnap, P; Javed, A; Liu, H.; Ozaip, S. (2019), Hate in the Machine: Anti-Black and Anti-Muslim Social Media Posts as Predictors of Offline Racially and Religiously Aggravated Crime, *British J. Criminol.*, 60, 93-117.
- Wyatt, M. and Silvester, J. (2015) Reflections on the labyrinth: Investigating black and minority ethnic leaders' career experiences, *Human Relations* 68(8): 1243-1269.
- Zempi, I. and Chakraborti, N. (2014) Islamophobia, Victimisation and the Veil, Basingstoke: Palgrave Macmillan