



The Need for New Tools to Break the Silos: Identity Categories in Hate Speech Legislation

Jen Neller¹

Birkbeck University of London, United Kingdom

Abstract

The implications of specifying certain identity categories have been widely debated in the context of hate crime laws and policies. However, they have been less thoroughly examined in the particular contexts of hate *speech*. Although the majority of laws regulating speech do not differentiate between identity categories, the ‘stirring up’ offences of the United Kingdom *Public Order Act 1986* are stratified along grounds of race, religion and sexual orientation. This article argues that, while the concerns raised about identity categories in relation to hate crime legislation are equally relevant to the stirring up provisions, the proposed solutions cannot automatically be transposed to hate speech offences. Accordingly, this article explores challenges that are encountered in attempts to make hate crime and hate speech legislation more inclusive before advancing some tentative suggestions for how hate speech laws might move beyond identity silos.

Keywords

Hate speech; hate crime; legislative reform; identity; vulnerability.

Please cite this article as:

Neller J (2018) The need for new tools to break the silos: Identity categories in hate speech legislation. *International Journal for Crime, Justice and Social Democracy* 7(2): 75-90. DOI: 10.5204/ijcjsd.v7i2.501.



This work is licensed under a [Creative Commons Attribution 4.0 Licence](https://creativecommons.org/licenses/by/4.0/). As an open access journal, articles are free to use, with proper attribution, in educational and other non-commercial settings. ISSN: 2202-8005

Introduction

According to Mason-Bish, 'The silo approach to hate crime policy is arguably one of its defining features and biggest problems' (2014: 167). What is referred to here as the 'silo approach' is the construction or enhancement of criminal offences penalising hatred, hostility or prejudice on grounds of a finite selection of discrete identity categories, such as race, religion, sexual orientation, disability or transgender identity. Both advocates and opponents of anti-hate legislation (for example, Schweppe 2012 and Heinze 2009, respectively) have argued that such an approach is too narrow to reflect the wide variety of experiences of hatred, and that the corresponding legislation is, therefore, partial and discriminatory. While the list of categories within a particular law may be expanded, such a list can never be all-inclusive nor account for the ways in which different types of identity might intersect, interact or be fluid.

These limitations have been discussed at some length in relation to hate crime laws (see below), but they have received relatively little attention in relation to hate speech legislation (exceptions include Bakalis 2017; Brown 2016; Heinze 2009). In the United Kingdom (UK), identity categories are present in two Acts pertaining to hate crime and one which covers certain forms of hate speech. The hate crime Acts are the *Crime and Disorder Act 1998*, which establishes a class of 'aggravated offences' for offences motivated by or demonstrating racial or religious hostility, and the *Criminal Justice Act 2003*, which requires hostility on grounds of race, religion, sexual orientation, disability or transgender identity to be taken into account during sentencing for an offence. This article focuses on the hate speech legislation, which is contained within Parts III and IIIA of the *Public Order Act 1986*. These provisions prohibit the stirring up of hatred on grounds of race, religion or sexual orientation. The inconsistency regarding the selection of identity categories protected under these three Acts is, therefore, plain to see. However, the extent to which this inconsistency in particular and the silo approach in general are problematic and, moreover, the most beneficial solutions remain highly contested.

The first section of this article briefly surveys the various criticisms that have been levelled at the silo approach. These criticisms are not viewed as signalling that hate speech legislation is, and can only be, counter-productive or harmful but are, instead, impetus for exploring potential avenues for improvement. The second section reviews ideas as to how more inclusive hate crime legislation might be formulated. Then, turning to hate speech legislation, the third section outlines the 'stirring up' offences that are found within the UK *Public Order Act 1986*. While the inclusion of discrete identity categories within this legislation can be subjected to the same criticisms as identity silos in hate crime laws, this article argues that hate speech presents distinct challenges to the construction of more inclusive legislation. These challenges are explored and some tentative solutions are suggested in the last section.

Hate speech legislation is often talked about as a type of hate crime legislation, but here I treat them as distinct so that I can compare them with minimal confusion. To refer to both, I use the term 'anti-hate' legislation.

The problems with identity categories

Lists of identity categories determine the scope of hate crime statutes around the world (Brown 2016). However, variation between these lists need not be perceived as a problem of inconsistency. As hate is a social phenomenon manifesting in different forms and against different groups in different and evolving contexts, attempts to combat it can be expected to vary in scope and content (Bleich 2011: 8). This section provides a brief overview and evaluation of criticisms that such legislative lists of identity categories have engendered.

Hierarchies of victims

The complex, flexible, intersectional and evolving nature of how we identify ourselves and how we are identified by others makes it impossible for the law to list all conceivable identity

categories. By naming certain grounds upon which the propagation of hatred is criminal, the implication is that hatred based on other grounds is, if not fully acceptable, at least less problematic (Aradau 2008: 297; Gerstenfeld 2013: 56; Hall 2005 in Schweppe 2012: 179; Heinze 2009). In the context of hate crime legislation, the resulting disparities in legal protection have been described as producing a 'hierarchy of victims' that contravenes principles of equality and universal rights (Brown 2016: 301; Hentoff 1993: 45; Jacobs and Potter 1998: 21; Mason 2014: 169; Moran 2014: 266; Schweppe 2012: 191). Additionally, in the context of hate speech, Heinze (2009) points to the risk of ever-increasing restrictions on expression where the recognition of one minority group as victimised leads to demands for the recognition of other groups, resulting in cumulative add-ons to the law.

In assessing these points, the distinction between identity categories and minority groups should be noted, as the former do not exclude individuals *per se*:

Hate crime legislation ... protects everyone who has a race, everyone who has a gender, and everyone who has a sexual orientation so long as they are attacked because of their status. In other words, hate crime laws punish attacks on everyone. (Bell 2002: 181)

Therefore, lists of protected categories do not formally compromise principles of equality. However, while all individuals might be included, identity category lists still prioritise certain *experiences* of hatred. For example, persons with disabilities might have the same protection as anyone else against racial hatred but, if the legislation does not encompass hatred on the grounds of disability, they will not receive protection against the forms of hatred that might most concern them. Furthermore, anti-hate legislation is generally framed with particular vulnerable groups in mind, especially when it is enacted as a result of the lobbying efforts of such groups (Jenness 1999). For example, the addition of sexual orientation to the UK stirring up hatred provisions was argued to be necessary due to evidence of anti-gay hate speech (see Goodall 2009: 214). Thus, even though the law technically encompasses hate speech against heterosexuals, in reality the law was enacted for the protection of homosexuals and bisexuals.

Essentialising and dividing

The construction of discrete, rigid classifications is inherently at odds with the unique individuals and experiences that they are supposed to represent. Identity categories are thus a manifestation of 'the law's propensity to categorise' (Grabham 2009: 186) and to 'deny multiplicity, complexity and ambiguity' (Moran and Sharpe 2004: 408; see also Berlant 1997: 19; Moran 2014: 268). Scholars have critiqued simplistic approaches that treat identity categories as 'either/or', or as simply the sum of separable and pre-determined categories (Grabham 2009; Lambie 2008; Moran and Sharpe 2004; Mason 2005). While different identity categories clumsily encapsulate different experiences, they also fail to account for the ways in which particular combinations of perceived identities may evoke particular prejudices and hostilities (Garland and Hodgkinson 2014; Walters and Brown 2016). Therefore, it should not automatically be assumed that different identities require the same protections from the law; greater nuance is required to remedy the 'gaps, silences and exclusions' (Moran 2007: 431) that are produced by current statutory formulations of identity.

Claims have also been made that, by focusing on specific identity categories, anti-hate laws have divisive effects within society and serve to entrench perceptions of difference. For example, Chakraborti and Garland (2012: 501) describe how 'approaching the issue of inclusion through the lens of group identity politics merely exacerbates existing problems, creating divisions among communities of identity rather than highlighting the shared nature of their victimization'. Jacobs and Potter (1998: 8) use the term 'Balkanisation' to describe this effect, arguing that identity politics encourage individuals to view themselves and others as members of competing groups.

This reflects the common perception that such legislation panders to the demands of specific groups rather than providing equal protections for everyone on specific grounds.

Aside from the negative connotations assigned to an entire region through this neologism, Jacobs and Potter's (1998) notion of Balkanisation can be criticised for failing to consider the ways in which individuals defy discrete and static identity silos and may not automatically internalise the limits or alterity of the tick-box identities with which they are presented. Furthermore, any approach that seeks to resolve disparities or animosities between groups will almost inevitably identify these groups in some way, and thereby become complicit in their existence as distinct conceptual and social entities. Yet to acknowledge difference is not necessarily to the detriment of social cohesion and can, instead, be an essential step towards genuine inclusivity. If a hate crime is a 'message crime' (Perry 2001: 10; see also Rosga 2001) (and hate speech even more overtly so), the message surely requires a response. 'Strategic essentialism' (Law 1996: 5) should also be recognised as having enabled activists to successfully lobby for protection against discrimination and disadvantage. Therefore, efforts to redress gaps and exclusions are at risk of being counterproductive if they negate differences in such a way that ultimately disempowers those who would speak against marginalisation.

Subjectification

A final angle of criticism, which follows on from essentialism and Balkanisation, stems from the production of subjects through identity categories and the implications that this has for individuals. Identity politics have been criticised for reinforcing problematic distinctions that classify everyone in relation to a singular biased standard (white, able-bodied, heterosexual, cisgender,² male, citizen, and so on) (Brown 1995: 61). Non-standard groups are therefore at risk of being defined negatively, as a lack or a deviance. Identity politics aim, then, to redefine the relationship between the centre and the periphery through concessions and special measures, rather than to bring the periphery *into* the centre (Fraser 1995: 82). Recognition of difference and disadvantage does not always amount to emancipation or equality, and demands for legal equality may be 'a feature of systematic injustice, not a remedy' (Spade 2011: 41).

Insofar as identity categories are necessarily essentialised, they represent abstract subjects; victims of certain forms of hatred are constructed as innocent 'damsels in distress' who the state can protect, save and emancipate, rather than complex, imperfect agents (Brown 1995: 106). The unavoidable discrepancy between the abstract stereotype and the unique reality results in what may be referred to as misinterpellation, where the subject who appears is not entirely the subject who was called forth by the law (Martel 2017). For example, the law's archetypal victim of sexual violence is young, female and modest, and the law's archetypal victim of homophobia is openly 'out' (Berg and Millbank 2009; Briddock 2016). Such preconceptions of victims can lead to the differential treatment of those who do not meet expectations: 'By predicating political strategies on innocent victimhood, violence against individuals who deviate from the ideal becomes less visible and more tolerable' (Lamble 2008: 27; see also Howard 1984; Moran 2014: 269; Moran and Sharpe 2004: 408).

Another effect of the interpellation of 'pure' victims is the extent to which persons who fit within certain categories may be portrayed as inherently vulnerable and, concurrently, how essentialised claims of inherent vulnerability can become central to demands for recognition and protection (Brown 1995; see also Jenness and Grattet 2001: 10; Sullivan 1999 in Walters 2014). For example, Jenness and Broad (1994: 417) note such investments in trauma in both feminist and in gay and lesbian activism, and Westbrook has examined how activist representations of transgender persons as always vulnerable to attack tell 'trans people that, even if they are not currently afraid of violence, they should be' (Westbrook 2008: 18). Law and activism that is premised on such inscriptions of vulnerability may be disempowering because such investment in vulnerability naturalises the systems that construct these groups as outsiders (Brown 1995;

see also Fraser 1995: 74) and encourages them to feel responsible for managing their own vulnerability.

Hate crime legislation beyond the silos

As noted above, activism around identity categories has, in many cases, proved to be a successful means of shedding light on inequality and obtaining redress. Therefore, there is a risk that the alternative to using identity categories might be the absence of any language with which to challenge marginalisation. Solutions that seek to overcome the limits and exclusions of identity categories, then, should ensure that no avenues of expression are closed off and no groups are inadvertently disempowered. The aim must be to open up possibilities (Butler 2007: viii) and move towards greater inclusivity.

In order for hate crime legislation to be meaningful, there must be a clear distinction between 'ordinary' crimes and those which should incur additional penalties. Thus, scholars seeking to decouple hate crime provisions from a finite list of categories have sought, instead, to define the scope of provisions according to the underlying reasons that make some categories but not others appropriate for inclusion. Such approaches are outlined and critiqued in this section.

Open-ended clauses and the community criterion

Perhaps the most straightforward solution proposed to the formal exclusion and inflexibility of listed identity categories is to indicate that the stipulated list is not exhaustive. This can take the form of an additional clause after the list that reads 'or any similar factor'. The benefits of such an approach are that it enables the law to be more flexible; it can respond to acts of hatred against unusual or unpopular groups that do not typically fit preconceived ideas about victims of hatred, enabling the law to carry the message that there is no group of persons against whom violence and degradation is acceptable. The advantage of this approach is, for example, that it would enable a crime perpetrated due to bias against scientologists to be prosecuted as a hate crime without a court having to determine whether Scientology constitutes a religion.

The parameters of an open-ended list of categories depend on the question of what constitutes 'similar' or 'any other such grounds'. Indeed, this may be linked to the idea that commensurate grounds deserve commensurate protection. However, where the categories that are listed include, for example, race, religion and sexual orientation discerning the commonalities between them according to which other categories might be deemed similar can be difficult. Even if race and sexual orientation are deemed immutable, religion is more widely perceived as a characteristic that is chosen. Alternatively, where it may be possible to conceal religious and sexual identities, race often cannot be hidden. If we consider all three categories to share a history of widespread violent persecution, however, what is deemed 'similar' can no longer encompass emerging groups or groups facing new forms of persecution (see Schweppe 2012: 190).

Attempting to find a common essence within any given set of stipulated categories may be fallacious. Instead, an 'other' category might only need to be deemed similar to only one listed category. This seems conceptually murky water, though. Should hatred against transgender identities be assumed to be commensurate with hatred premised on sexual orientation? On what basis can disability be argued to be a similar category to any other? Can homelessness, welfare status, drug addiction or support of a particular football team be described as sufficiently similar to categories that are explicitly protected? *Should* they be?

Schweppe (2012) argues that one of the main benefits of the open-ended approach is that it does not require a normative assessment. For Schweppe, the requirement for protection should be that an attack was motivated by hostility towards particular personal characteristics that the victim shares with an 'identifiable social group' or a community. This, she argues, 'places the concerns of the discipline of hate studies at its core: that is, the desire to combat the dehumanisation of the

other, while simultaneously being inclusive and nondiscriminatory' (Schweppe 2012: 191). Indeed, Schweppe expressly rejects consideration of the extent to which the category in question is commensurate with a listed category (2012: 190). Similarity is, therefore, granted a wide interpretation but it is not necessarily clear what constitutes a community. It is relatively common these days for a gay, LGBT (lesbian, gay, bisexual and transgender) or queer community to be spoken of, but can persons with disabilities be said to form a community? What if we consider gender as a ground that motivates acts of hatred (Mason-Bish 2014)? Can women be said to comprise a community? And what then about the 'heterosexual community' or the 'male community'? Can such communities be said to exist in a meaningful way? The focus on communities seems to engender a shift from categories to groups. Thus, instead of a hate crime charge being brought because an attack was motivated by racial hatred, it would be brought because it was motivated by hatred of a particular racial community. This is not necessarily a criticism as it is likely to be more faithful to the harms that the lobbyists and law-makers intended the legislation to address, but the difficulty of defining a community would have to be attended to.

Additionally, issues relating to intersectionality and essentialism are not resolved by the community criterion, as individuals are not afforded protection unless they are perceived as belonging to a group. For example, Sophie Lancaster, who was fatally assaulted for belonging to the Goth subculture (see Garland 2010; Garland and Hodkinson 2014), would be defined as a victim of hate crime under the community criterion, but an eccentric character who does not fit any such singular cultural label and who is attacked because of their difference would not be defined as such.

Mason (2014) argues that open-ended clauses regarding the scope of hate crime laws can render the application of hate crime protections too wide. In her examination of two Australian cases where child sex offenders were ruled to have been victims of hate crime, Mason considers the symbolic effects of hate crime law, arguing that its application to child sex offenders signalled that this group is of comparable value to society as minority races, religions or sexualities. For Mason, these cases demonstrated 'that we need more concrete guidance on the kinds of difference that warrant protection' (2014: 173). Schweppe's community criterion does not seem helpful here: if we wish to include persons with disabilities, for example, but exclude child sex offenders, can the former group really be said to face attacks on the basis of belonging to a community more than the latter?

Perceived vulnerability and difference

Chakraborti and Garland (2012) suggest that the language of hatred itself might be unhelpful and might serve to exclude many people from protection. Thus, rather than legislating against hate crimes, Chakraborti and Garland call for legislation against a broader concept of bias crimes which encompasses all situations where the victim is targeted as a result of their perceived vulnerability or 'difference'. This builds upon Perry's (2001) definition of hate crime as a process of inscribing difference and asserting the structural inferiority of a group. Chakraborti and Garland argue that such a definition fails to account for instances of targeted violence that are more opportunist and/or more individualised, as well as instances where 'traditional' power dynamics are inverted and persons are targeted for their membership of a supposedly more dominant social group (2012: 504-505). Thus, '[c]onceiving of these offences purely as a mechanism of oppression or subordination overplays what for some perpetrators will be an act borne from boredom, jealousy, convenience or unfamiliarity with "difference"' (2012: 506). Chakraborti and Garland therefore propose a nuanced understanding of how a person's identity characteristics can combine with their particular situational contexts to render them vulnerable in the eyes of a perpetrator.

The ability of Chakraborti and Garland's (2012) approach to account for intersectionality is one of its main benefits. Vulnerability is not uniformly attached to discrete and self-evident identity categories but is perceived as arising through a particular combination of factors that could be unique to the victim or particular to the moment in question. For example, this enables gender to be taken into consideration without either automatically conflating femaleness with vulnerability or assessing the extent to which the victim complies with pre-determined ideas about what femaleness is or should be (Stanko 2001: 318). Another advantage of the vulnerability approach is that an individual need not be recognised as a member of any community in order to be recognised as a victim and, therefore, protection is not differentiated according to the visibility of particular groups (Chakraborti and Garland 2012: 509). Additionally, this approach allows for vulnerability and hostility towards difference to be considered in tandem. Walters (2013) argues that this is necessary in relation to a case of organised sexual abuse of white girls by Asian men, for example, where both the vulnerability of the victims as young females and their different race should be considered in relation to the perpetrators' motivations. The combination of vulnerability and hostility to difference is also increasingly recognised as relevant in instances of crime against persons with disabilities (Garland 2010).

However, Chakraborti's and Garland's (2012) approach produces the challenge of how to assess a person's actual or perceived vulnerability and of whether victims can be rendered not vulnerable or different enough. For example, difficulty surrounding the notion of vulnerability was demonstrated in a recent UK case where a judge initially granted a suspended sentence to a physically abusive husband on the basis that his wife was too intelligent and well-connected to be considered vulnerable to his abuse (James 2017). To take another example, aggression against feminists might not be precipitated by their vulnerability so much as by their perceived strength and influence. Here, the attack seems more of an attempt to weaken and inscribe vulnerability upon the victim rather than a response to already-existing vulnerability. Rather than vulnerability, such an attack may be considered on the basis of difference but, as the label of 'feminist' has come to be adopted by a wider range of people, the question is difference to what or whom?

Despite their weaknesses, these approaches are path-breaking explorations that lead away from rigid identity categories and towards greater inclusivity. In the final section of this article, these approaches are used to highlight the differing set of challenges that arise in relation to hate speech, moving towards some suggestions on ways forward. First, however, I briefly outline the law to which this analysis will be anchored.

The offence of stirring up hatred in the UK

The *Public Order Act 1986* (POA) introduced several provisions that broadly fall within contemporary understandings of 'hate speech'. Section 4 prohibits the use of threatening, abusive or insulting words or behaviour that are either intended or likely to cause another person to fear that violence will be used against them or to provoke them to use violence themselves. Section 4A prohibits the use of such language or behaviour that *causes* another person harassment, alarm or distress. Section 5 prohibits the use of threatening or abusive words or behaviour that is *likely* to cause another person harassment, alarm or distress.

The POA also includes a reformulation of the incitement to racial hatred provisions that were first introduced in the *Race Relations Act 1965*. These provisions now form Part III of the POA and they begin as follows:

- (1) A person who uses threatening, abusive or insulting words or behaviour, or displays any written material which is threatening, abusive or insulting, is guilty of an offence if—
 - (a) he intends thereby to stir up racial hatred, or

(b) having regard to all the circumstances racial hatred is likely to be stirred up thereby.

Part IIIA pertaining to the stirring up of religious hatred was introduced into the POA by the *Racial and Religious Hatred Act 2006*. The wording of this Part was subsequently amended to include stirring up hatred on the grounds of sexual orientation by the *Criminal Justice and Immigration Act 2008*. However, there is a higher threshold for the offences of stirring up hatred on grounds of religion or sexual orientation than for stirring up racial hatred: the words or behaviour must be threatening (not just abusive or insulting) and must be intended to stir up hatred (not just likely to do so).³ Part IIIA also includes controversial free speech provisions pertaining to religious belief and sexual orientation that were not found necessary for racial hatred (Sandberg and Doe 2008: 985; Vanderbeck and Johnson 2011).

The stirring up provisions are distinct from the other speech offences contained in the POA in four ways. Firstly, they address the speech itself as a problem rather than viewing it as a problem only in relation to its potential consequences (such as the distress of a listener/viewer). Secondly, they include speech that is *not* addressed to those whom it disparages, but that instead seeks to spread hatred within a group. These two points are significant because they mean that the offence covers situations where there are no identifiable victims.

Thirdly, as a result of the absence of individual victims, the stirring up provisions can only deal with hatred against groups. To be intelligible to the law, these projected victim categories are necessarily essentialised preconceptions of the sorts of people whom the law wants to protect. Indeed, in light of Zedner's (2007) analysis of the legal shift from backward-looking reactions to crime to forward-looking predictions of crime—and thus the increasing phenomenon of 'pre-crime'—these groups might be conceived of as 'pre-victims'. The decision whether to prosecute under the stirring up provisions therefore amounts to an assessment of whether the speech is intended (or likely) to incite hatred against an imagined community of pre-victims.

Fourthly and correspondingly, the stirring up offences are the only speech-related provisions that refer to identity categories. Other provisions, such as Section 5, do encompass acts that may be prosecuted as aggravated on the grounds of race, religion or sexual orientation; however, threatening or abusive language that is used to cause harassment, alarm or distress is an offence under Section 5 regardless of whether it is also classified as, for example, racist. Conversely, the stirring up of hatred is *only* an offence in relation to one of the three established identity categories.

Hate speech legislation: Breaking the silos

In the search for alternatives to rigid identity categories, the stirring up provisions present different challenges to those discussed above in relation to hate crime laws. Firstly, hate speech legislation is not constrained by the need to distinguish between 'ordinary' crimes and hate crimes. Whereas hate crime provisions enhance the penalties that are applied to particular pre-existing offences, Parts III and IIIA of the POA create new offences that would not otherwise be prosecutable. However, the normative question of who should and should not be protected remains. Secondly, the concern to avoid essentialising victims is also more complicated in relation to the stirring up offences due to the absence of actual victims. Thirdly, such legislation must be reconciled with freedom of speech protections. Each of these issues is explored in this section.

The absence of victims

As noted above, the stirring up provisions do not require speech to cause, or even to be likely to cause, harm to individuals but is, instead, premised on the harm that hate speech *could* cause to a community or to inter-community relations. Such harms might include fear, discrimination, harassment, hostility, vandalism or violence, all of which would disadvantage the targeted

community and be detrimental to society overall. Schweppe's model of replacing specified identity categories with the broader criterion of a community under attack could, therefore, be an appropriate approach to making the stirring up provisions more inclusive. This is especially appealing as the criticism that this approach excludes those who fall in between communities does not apply: the stirring up provisions *can only be applied to groups*, as anything that is directed towards an individual would come under a different area of law.

Furthermore, questions about what can be considered a community and how to account for intersectionality are less problematic in relation to hate speech, as the target community will be interpellated by the speech in question. For example, regardless of whether the hate speech targets all 'queers', lesbians only or butch lesbians only, for the purpose of a prosecution, the speech itself defines the victim community without any need for the inherent differentness or vulnerability of the group to be inscribed in law. It may be possible, therefore, to have legislation that prosecutes hate speech directed against any community, as that community is defined within the speech itself. The likely criticism of such an approach is that it is too broad and amounts to an unacceptable imposition on freedom of speech; this issue is examined below. This approach also falls prey to a hate speech equivalent of Mason's (2014) scenario, whereby the prosecution of hate speech against child sex offenders, a group demarcated by hate speech itself, would be enabled on the basis that child sex offenders may subsequently be marginalised and disadvantaged.

While the absence of manifest victims can be seen as helpful to Schweppe's (2012) community-based approach, it is far more problematic for Chakraborti and Garland's (2012) vulnerability approach. This is because the vulnerability approach unavoidably relies upon the presence of a victim whose vulnerability, or perceived vulnerability, can be assessed. An assessment of vulnerability under the stirring up offences would instead require the appraisal of an abstract and essentialised group, and would subsequently have to rely on stereotypes and assumptions.

Who or what to protect?

One challenge with hate crime law is that applying penalty-enhancing provisions too widely undermines the message that such laws communicate about the depravity of certain motivations or biases. Similar concerns have been voiced in relation to hate speech, where applying such provisions to protect everyone equally risks negating the significance of difference, especially in relation to the experiences of historically oppressed groups. Thus, Brown asks how hate speech can be restricted:

... in the name of equality ... without making all persons equally viable as victims of such speech, thereby forfeiting a political analysis that recognises the specific function of hateful epithet in sustaining the subordination of historically subordinated peoples. (Brown 2002: 424)

We seem, then, to be trapped between essentialising and negating difference, in addition to again sliding into an approach that focuses on groups rather than categories.

An answer to Brown's question lies in the range of contextual factors that could be considered. In some instances, consideration of historical subordination may be crucial to understanding the intentions of the speaker and the likely effects of the speech. However, this should not prejudice consideration of speech that is aimed at forging new divisions in society rather than usurping well-defined existing ones. This perhaps echoes the concept of 'othering' that is developed by Perry and the significance that Chakraborti and Garland place on notions of 'difference'; fear of difference and the impulse to assert superiority over it may be just as harmful when it is a newly constructed distinction as when it is an all-too-familiar one. Inclusivity and recognition of historical subordination are not necessarily mutually exclusive so long as historical factors are

only one criterion among others that are considered in the assessment of whether hatred was intended or likely to be stirred up.

Freedom of speech

For some commentators, all hate speech laws are inimical to freedom of speech and hate speech is best combatted through the freedom to reply within a 'free marketplace' of ideas (see Haiman 1993; Heinze 2006: 553). However, others argue that this formal freedom obscures power imbalances whereby equal freedom to speak does not equate to equal consequences for doing so or an equal ability to be heard (Lillian 2007: 735; MacKinnon 1996: 102). Moreover, Lawrence (1993: 68) has argued that the purpose of hate speech is to exacerbate such disparities and to shrink the space in which the target can receive a fair hearing. The challenges for hate speech legislation, then, are to level this playing field without creating or exacerbating gaps or silences.

A further argument is that, to be more inclusive, more speech would need to fall within the purview of the law. Any expansion of the current provisions may, therefore, be resisted on the grounds that this would contribute to a trend of increasingly restrictive censorship of speech (Heinze 2009). Furthermore, prosecutions under the stirring up provisions have tended to punish relatively powerless outsiders for isolated instances of hate speech and thus to have amounted to the punishment, rather than protection, of marginalised persons (European Commission against Racism and Intolerance 2015: 5; Malik 2009: 105; see also Spade 2011: 64-65).

However, greater inclusivity within the stirring up provisions need not necessarily result in a broad infringement on freedom of speech for three reasons. Firstly, if freedom of speech does not grant a license to stir up hatred on the grounds of race, religion or sexual orientation, suggesting that it does contain a right to stir up hatred on other grounds seems arbitrary and illogical. Therefore, more inclusive legislation would not necessarily result in a *further* restriction of free speech; rather, neither the current provisions nor any potentially more inclusive iterations would *necessarily* impede free speech at all, as it is legally delineated (Committee on the Elimination of Racial Discrimination 2013: para 45).

Secondly, the stirring up provisions have not been widely used. The expansion of the stirring up offences to include hatred on the grounds of religion and sexual orientation has not led to a slew of prosecutions. Indeed, the infrequency with which they are invoked has been levelled as a criticism against them and as an argument for the futility of extending them to cover other categories (UK Law Commission 2014). While such an argument seems to overlook the extent to which such provisions may act as a deterrent and to ignore their symbolic value, the fear of a slippery slope towards massively restricted and/or chilled speech (Edwin Baker 2009: 154; Heinze 2009) also appears unfounded as the stirring up provisions represent an incremental and closely monitored development (Bleich 2011: 6; Rumney 2003: 139).

Thirdly, it may be possible to balance the widening of the legislation in relation to identity categories by narrowing the legislation in other respects. Consideration of a range of contextual factors could be used to create a stricter test of whether prosecution is in the public interest. The six-point test developed by Article 19 (2012), an international non-government organisation, and adopted in the United Nations High Commissioner for Human Rights' (2013) Rabat Plan of Action may provide a useful guide here for establishing a threshold of severity and influence. These six points can be summarised as the social and political context of the expression; the status of the speaker; the intent of the speaker; the content of the expression; the extent and magnitude of the expression; and the likelihood and imminence of harm occurring as a result. While these points would likely require some further refinement for use in domestic legislation, such a means of establishing a nuanced and flexible threshold would help to ensure that prosecutions are limited to serious and consequential cases.

Even with the above three arguments taken into consideration, inclusion of *all* grounds of hatred within the scope of the law would still not be acceptable. In particular, the expression and incitement of hatred towards political groups must be permissible within a functioning democracy. This does not include personal threats, harassment or other criminal acts but the expression of disgust, hostility, derision and perceptions of inferiority towards people because of the political views they hold must be possible. A person's politics are not an inconsequential private matter but, rather, can be representative of how they view those around them and can ultimately influence issues of governance that affect other people in profound ways. As such, an expression of hostility towards people who vote a certain way or hold a certain political belief can be an expression of a judgement rather than of prejudice, and to prohibit such expressions would be, *inter alia*, to prohibit expressions of hatred against a group for their oppressive actions. This would seem to cross the line into censorship and could exacerbate inequalities rather than aid in their redress.

So might a specific exemption protecting freedom of political speech be in order? I argue against this for two reasons. Firstly, freedom of speech should be the default and speech which is prohibited should be narrowly circumscribed. Secondly, there will inevitably be some difficulty in determining what constitutes political speech or a political group. So, as with the distinction between hate crimes and 'ordinary' crimes, the criteria that are used to distinguish expressions of hatred that it is desirable to criminalise from expressions of hatred against political targets that it is desirable to protect should be established as objectively as possible. This should not be a question of whom to protect but, rather, of determining precisely that from which it is desirable to protect people and society.

I suggest that a distinction between 'pertinent' and 'arbitrary' characteristics could be used to determine whether speech denigrates a group based upon relevant or irrelevant factors. Relevant and permissible factors might include a group's subscription to certain political beliefs, their active participation in a particular event or their behaviour that in some significant way affects public life. Arbitrary characteristics, alternatively, are those that have no relevance to public life and/or to which sweeping negative connotations can only be ascribed through prejudice and stereotyping. Thus, in the discussion of a political party, insulting members of the party collectively for the policies held and actions taken by the group would be acceptable. However, it would be arbitrary to then state that the skin colour of a number of party members was the root of the problem; such factors have no objective bearing on political activities, and to claim that people who look a certain way will behave a certain way is to pre-judge them.

To take another example that has been discussed in this article, a concept of arbitrary characteristics would not extend protection to child sex offenders. The stirring up of hatred against child sex offenders incites hatred against a group of people because of illegal actions that they have performed and of which they have been found guilty in a court of law. Inciting the commission of crimes against such persons would be criminal but stating that they are an evil scourge upon society, for example, would not.

The example of inciting hatred against feminists is more difficult. Such hatred is ostensibly targeted towards political beliefs rather than particular types of people but it can also be seen as an attempt to shut down certain voices and, thereby, to perpetuate historically entrenched oppression. Therefore, the concept of arbitrary characteristics needs to consider the effects of the speech as well as its content. Expressing hatred towards feminists because they are 'anti-men', because they are aggressive or because 'it is women who are privileged in society' should be allowed, not only in deference to the freedom of speech of the speaker but also so that counter arguments can be made and open discussion can occur. Expressing hatred towards feminists because women do not deserve equality and/or do not deserve a voice, however, is inimical to open discussion and is ultimately stirring up hatred based on stereotypes of women, rather than reacting to the beliefs of feminists. A person's gender has no inherent bearing on their beliefs,

activities or capacities, so such expressions of hatred can be proscribed on the grounds that they target an arbitrary characteristic.

Other difficult examples include hatred against homosexuals, sex workers and drug addicts, as membership of each of these groups tend to be defined by the actions taken. One way to deal with this would be to imbue the notion of arbitrary characteristics with a division between public and private characteristics. Thus, the sexual orientation of an individual might be regarded as a characteristic (see Berlant 1997) that should have no bearing on public life, so it is impermissible to incite hatred on such grounds. Voicing disapproval of homosexual activity would be permissible, for example, but not inciting hatred against consenting adults who engage in such activity in private. This might be a fine line to draw but it would place no more discretion in the hands of a court than the current legislation. The case of sex workers and addicts is less straightforward as there is perhaps more of a case to be made that their activities impact society and public life, even though they cannot be said to involve public activities *per se*. Incitement to hatred of these groups, therefore, might not be encompassed within this model. Such speech would certainly not be helpful, but it would not be criminal unless it crossed the threshold of inciting violence.

Perhaps the exclusion of such groups is a flaw in the arbitrary characteristics approach; or perhaps it is not, due to the rarity of incitement to hatred on these grounds and the need for open conversations around these topics. Indeed, these examples may indicate that the protected grounds can—and maybe even should—vary between hate crime and hate speech legislation due to the different issues surrounding these distinct areas of law. This is perhaps exemplified by the example of hatred on political grounds. On the one hand, expressions of hatred and derogatory terms for persons adhering to one or another political faction are rampant in the current political climate in the UK. For example, necklaces displaying the words ‘Fuck the Tories’ can be purchased from the website www.fuckthetories.co.uk, which greets you with the phrase ‘Welcome Tory hater’. Such incitement to hatred, while possibly causing offence to some, is generally tolerated within the democratic competition of ideologies. On the other hand, in line with Schedule 21 of the *Criminal Justice Act 2003*, the murder of UK Member of Parliament Jo Cox was described by the judge who sentenced her killer as exceptionally serious due to the attacker’s political, racial and ideological motivation (Wilkie 2016). These examples suggest that political hatred is a category that is more appropriate for specific legal redress in relation to criminal motivation than to speech.

Conclusion

This article explores the challenges associated with attempts to divorce anti-hate legislation from finite lists of identity categories that fail to represent the lived realities of the law’s subjects. Examining the nuances of this task has revealed that there are significant differences between the challenges that arise in relation to hate crime and those that emerge in relation to hate speech. For hate crime laws, the primary challenge lies in the need to distinguish hate crimes from ‘ordinary’ crimes whereas, for hate speech laws, the difficulty lies in the distinction between hate speech and speech that should be safeguarded by freedom of expression protections. Additionally, hate crime laws face the challenge of responding to acts that *have* affected unique *individuals* in unique ways, whereas the stirring up offences under the UK *Public Order Act 1986* are premised only on the *potential* effects that may be experienced by certain *types* of people.

The analysis in this article suggests that the stirring up hatred provisions could be decoupled from static lists of identity categories if an offence is instead determined by three conditions:

1. Hatred is stirred up against a discernible ‘community’ or group of persons;

2. The hatred is directed at the group on the grounds of an arbitrary characteristic, that is, a characteristic that a) has no direct bearing on the views or activities of the group's members or b) is fundamentally a private matter that does not impact public life; and
3. A threshold of severity and influence of the speech is attained.

These three factors might enable hate speech legislation to be implemented in a more inclusive and flexible manner yet safeguard political debate, avoid the prosecution of marginal characters and defy the supposed slippery slope of censorship. Critics will no doubt find the notion of arbitrary characteristics too vague to be practicable, a criticism that is levelled at one or another aspect of most hate speech laws. But it could also be argued that such vagueness is an advantage: it allows for the specific circumstances of a case to be meaningfully considered; it allows for unforeseen forms of hatred to be challenged; and it allows for guidance to be provided through case law as situations emerge rather than being dictated in the legislation at the time of enactment.

This article does not hold all of the solutions to the challenges discussed here. Indeed, this article may make a more convincing case that there are no universal or consensual solutions to be found. However, it has illustrated that solutions proposed with hate crime in mind cannot be assumed to be appropriate for hate speech. Closer attention to the stirring up provisions reveals that they require separate consideration and dedicated solutions to effectively overcome the limitations inherent to all laws that use identity silos to determine who is, and who is not, a victim.

Correspondence: Jen Neller, Associate Tutor and Doctoral candidate, School of Law, Birkbeck University of London, Malet Street, Bloomsbury, London WC1E 7HX, United Kingdom. Email: jenkarenh@gmail.com

¹ The author thanks Professor Les Moran, attendees of the International Network of Hate Speech 2016 Conference in Limerick and an anonymous reviewer for their helpful comments on earlier versions of this article.

² 'Cisgender' is an antonym of 'transgender'. The term describes persons whose gender identification aligns with the sex they were attributed at birth.

³ This is especially significant considering that hate speech against Jews can be prosecuted as stirring up racial hatred while hate speech against Muslims can only be prosecuted as stirring up hatred on the grounds of religion, meaning that convictions are easier to secure for the protection of one group than the other (Brown 2016: 291).

References

- Aradau C (2008) Forget equality? Security and liberty in the 'War on Terror'. *Alternatives: Global, Local, Political* 33(3): 293-314. DOI: 10.1177/030437540803300302.
- Article 19 (2012) *Prohibiting Incitement to Discrimination, Hostility or Violence*. London: Article 19.
- Bakalis C (2017) The victims of hate crime and the principles of the criminal law. *Legal Studies*. Epub ahead of print 5 June 2017. DOI: 10.1111/lest.12171.
- Bell J (2002) *Policing Hatred: Law Enforcement, Civil Rights, and Hate Crime*. New York: New York University Press.

- Berg L and Millbank J (2009) Constructing the personal narratives of lesbian, gay and bisexual asylum claimants. *Journal of Refugee Studies* 22(2): 195-223. DOI: 10.1093/jrs/sep010.
- Berlant L (1997) *The Queen of America goes to Washington City*. Durham, North Carolina: Duke University Press.
- Bleich E (2011) *The Freedom to be Racist? How the United States and Europe Struggle to Preserve Freedom and Combat Racism*. Oxford: Oxford University Press.
- Briddock A (2016) The recognition of refugees based on sexual orientation and gender identity in the UK: An overview of law and procedure. *Birkbeck Law Review* 4(1): 123-157.
- Brown A (2016) The 'who?' question in the hate speech debate: Part 1. *Canadian Journal of Law and Jurisprudence* 29(2): 275-320. DOI: 10.1017/cjlj.2016.13.
- Brown W (1995) *States of Injury: Power and Freedom in Late Modernity*. Princeton, New Jersey: Princeton University Press.
- Brown W (2002) Suffering the Paradoxes of Rights. In Brown W and Halley J (eds) *Left Legalism/Left Critique*. Durham, North Carolina: Duke University Press.
- Butler J (2007 [1990]) *Gender Trouble: Feminism and the Subversion of Identity*. London: Routledge.
- Chakraborti N and Garland J (2012) Reconceptualising hate crime victimization through the lens of vulnerability and 'difference'. *Theoretical Criminology* 16(4): 499-514. DOI: 10.1177/1362480612439432.
- Committee on the Elimination of Racial Discrimination (CERD) (2013) *General Recommendation No. 35: Combating Racist Hate Speech*. UN Doc: CERD/C/GC/35.
- Edwin Baker C (2009) Autonomy and hate speech. In Hare I and Weinstein J (eds) *Extreme Speech and Democracy*: 139-157. Oxford: Oxford University Press.
- European Commission against Racism and Intolerance (ECRI) (2015) *ECRI General Policy Recommendation No. 15 on Combating Hate Speech*. Available at http://www.coe.int/t/dghl/monitoring/ecri/activities/GPR/EN/Recommendation_N15/RE C-15-2016-015-ENG.pdf (accessed 10 May 2016).
- Fraser N (1995) From redistribution to recognition? Dilemmas of justice in a 'post-socialist' age. *New Left Review* 212: 68-93.
- Garland J (2010) 'It's a mosher just been banged for no reason': Assessing targeted violence against goths and the parameters of hate crime. *International Review of Victimology* 17(2): 159-177. DOI: 10.1177/026975801001700202.
- Garland J and Hodkinson P (2014) 'F**king freak! What the hell do you think you look like?': Experiences of targeted victimization among Goths and developing notions of hate crime. *British Journal of Criminology* 54(4): 613-631. DOI: 10.1093/bjc/azu018.
- Gerstenfeld P (2013) *Hate Crimes: Causes, Controls, and Controversies*. London: SAGE.
- Goodall K (2009) Challenging hate speech: Incitement to hatred on grounds of sexual orientation in England, Wales and Northern Ireland. *International Journal of Human Rights* 13(2-3): 211-232. DOI: 10.1080/13642980902758127.
- Grabham E (2009) Intersectionality: Traumatic impressions. In Grabham E, Cooper D, Krishnadas J and Herman D (eds) *Intersectionality and Beyond: Law, Power and the Politics of Location*: 183-201. Abingdon, Oxon: Routledge-Cavendish.
- Haiman F (1993) *'Speech Acts' and the First Amendment*. Carbondale, Illinois: Southern Illinois University Press.
- Heinze E (2006) Viewpoint absolutism and hate speech. *Modern Law Review* 69(4) 543-582. DOI: 10.1111/j.1468-2230.2006.00599.x.
- Heinze E (2009) Cumulative jurisprudence and minority rights: The example of sexual minorities and hate speech. *International Journal of Human Rights* 13(2-3): 193-209. DOI: 10.1080/13642980902758119.

- Hentoff N (1993) Hate crimes, should they carry enhanced penalties? No: Equality among victims. *ABA Journal* 79(May): 44-45.
- Howard J (1984) The 'normal' victim: The effects of gender stereotypes on reactions to victims. *Social Psychology Quarterly* 47(3): 270-281. DOI: 10.2307/3033824.
- Jacobs J and Potter K (1998) *Hate Crimes: Criminal Law and Identity Politics*. Oxford: Oxford University Press.
- James E (2017) Cricketer who beat wife with a bat and forced her to drink bleach is spared jail. *Manchester Evening News*, 27 March. Available at <http://www.manchestereveningnews.co.uk/news/greater-manchester-news/cricketer-who-beat-wife-bat-12802130> (accessed 24 September 2017).
- Jenness V (1999) Managing differences and making legislation: Social movements and the racialization, sexualization, and gendering of federal hate crime law in the US, 1985-1998. *Social Problems* 46(4): 548-571. DOI: 10.2307/3097075.
- Jenness V and Broad K (1994) Antiviolence Activism and the (in)visibility of gender in the gay/lesbian and women's movements. *Gender and Society* 8(3): 402-423. DOI: 10.1177/089124394008003007.
- Jenness V and Grattet R (2001) *Making Hate a Crime: From Social Movement to Law Enforcement*. New York: Russell Sage Foundation.
- Lamble S (2008) Retelling racialized violence, remaking white innocence: The politics of interlocking oppressions in transgender day of remembrance. *Sexuality Research and Social Policy* 5(1): 24-42. DOI: 10.1525/srsp.2008.5.1.24.
- Law I (1996) *Racism, Ethnicity and Social Policy*. New York: Prentice Hall.
- Lawrence C (1993) If he hollers let him go: Regulating racist speech on campus. In Matsuda M, Lawrence C, Delgado R and Williams Crenshaw K (eds) *Words that Wound: Critical Race Theory, Assaultive Speech and the First Amendment*: 53-88. Boulder, Colorado: Westview Press.
- Lillian D (2007) A thorn by any other name: Sexist discourse as hate speech. *Discourse and Society* 18(6): 719-740. DOI: 10.1177/0957926507082193.
- MacKinnon C (1996) *Only Words*. Cambridge, Massachusetts: Harvard University Press.
- Malik M (2009) Extreme speech and liberalism. In Hare I and Weinstein J (eds) *Extreme Speech and Democracy*: 96-122. Oxford: Oxford University Press.
- Martel J (2017) *The Misinterpellated Subject*. Durham, North Carolina: Duke University Press.
- Mason G (2005) Hate crime and the image of the stranger. *British Journal of Criminology* 45(6): 837-859. DOI: 10.1093/bjc/azi016.
- Mason G (2014) Victim attributes in hate crime law. *British Journal of Criminology* 54(2): 161-179. DOI: 10.1093/bjc/azt073.
- Mason-Bish H (2014) We need to talk about women: Examining the place of gender in hate crime policy. In Chakraborti N and Garland J (eds) *Responding to Hate Crime*: 169-182. Bristol, United Kingdom: Policy Press.
- Moran L (2007) Invisible minorities: Challenging community and neighbourhood models of policing. *Criminology and Criminal Justice* 7(4): 417-441. DOI: 10.1177/1748895807082065.
- Moran L (2014) LGBT hate crime. In Hall N, Corb A, Giannasi P and Grieve JGD (eds) *The Routledge International Handbook on Hate Crime*: 266-277. New York: Routledge.
- Moran L and Sharpe A (2004) Violence, identity and policing: The case of violence against transgender people. *Criminal Justice* 4(4): 395-417. DOI: 10.1177/1466802504048656.
- Perry B (2001) *In the Name of Hate: Understanding Hate Crimes*. New York City, New York: Routledge.
- Rosga A (2001) Deadly words: State power and the entanglement of speech and violence in hate crime. *Law and Critique* 12(3): 223-252. DOI: 10.1023/A:1013788321729.

- Rumney P (2003) The British experience of racist hate speech regulation: A lesson for First Amendment absolutists? *Common Law World Review* 32(2): 117-160.
- Sandberg R and Doe N (2008) The strange death of blasphemy. *The Modern Law Review* 71(6): 971-986. DOI: 10.1111/j.1468-2230.2008.00723.x
- Schweppe J (2012) Defining characteristics and politicising victims: A legal perspective. *Journal of Hate Studies* 10(1): 173-198.
- Spade D (2011) *Normal Life: Administrative Violence, Critical Trans Politics, and the Limits of Law*. Brooklyn, New York: South End Press.
- Stanko E (2001) Re-conceptualising the policing of hatred: Confessions and worrying dilemmas of a consultant. *Law and Critique* 12(3): 309-329. DOI: 10.1023/A:1013784203982.
- UK Law Commission (2014) *Hate Crime: Should the Current Offences Be Extended?* London: UK Government.
- United Nations High Commissioner for Human Rights (2013) Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence. UN Doc: A/HRC/22/17/Add.4. Available at http://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf (accessed 1 March 2018).
- Vanderbeck R and Johnson P (2011) 'If a charge was brought against a saintly religious leader whose intention was to save souls...': An analysis of UK parliamentary debates over incitement to hatred on the grounds of sexual orientation. *Parliamentary Affairs* 66(4): 652-673. DOI: 10.1093/pa/gsr022.
- Walters M (2013) Why the Rochdale Gang should have been sentenced as 'hate crime' offenders. *Criminal Law Review* 2: 131-144.
- Walters M (2014) Conceptualizing 'hostility' for hate crime law: Minding 'the minutiae' when interpreting section 28(1)(a) of the Crime and Disorder Act 1998. *Oxford Journal of Legal Studies* 34(1): 47-74. DOI: 10.1093/ojls/gqt021.
- Walters M and Brown R (2016) Causes and motivations of hate crime. *Equality and Human Rights Commission Research Report* 102.
- Westbrook L (2008) Vulnerable subjecthood: The risks and benefits of the struggle for hate crime legislation. *Berkeley Journal of Sociology* 52: 3-23.
- Wilkie A (2016) Sentencing remarks of Mr Justice Wilkie: R v Mair (Jo Cox murder). Available at <https://www.judiciary.gov.uk/judgments/sentencing-remarks-of-mr-justice-wilkie-r-v-mair-jo-cox-murder/> (accessed 24 September 2017).
- Zedner L (2007) Pre-crime and post-criminology? *Theoretical Criminology* 11(2): 261-281. DOI: 10.1177/1362480607075851.

Legislation cited

- Crime and Disorder Act 1998* (UK).
- Criminal Justice Act 2003* (UK).
- Criminal Justice and Immigration Act 2008* (UK).
- Public Order Act 1986* (UK).
- Race Relations Act 1965* (UK).
- Racial and Religious Hatred Act 2006* (UK).