

The “Who?” Question in the Hate Speech Debate:

Part 2: Functional and Democratic Approaches

Alexander Brown

Part 1 of this two-part article introduced the “Who?” question in the hate speech debate as well as this meta-question: on the basis of which moral and practical considerations should authorities specify the proper scope of incitement to hatred laws.¹ Part 1 looked at *consistency*, *practical* and *formal* approaches. Here in Part 2 I shall focus instead on *functional* and *democratic* approaches. Along the way I shall also assess a range of substantive arguments about which particular characteristics should and should not be protected by incitement to hatred laws. My main conclusion shall be that each of the approaches has its strengths and weaknesses and that, partly because of this, no single approach is adequate *by itself* as a tool for specifying the proper scope of incitement to hatred laws, but also, by the same token, no approach should be ruled out entirely. Instead, the best strategy is one that combines together all five approaches in reasonable compromise given the law, the characteristic and the context. At the end of Part 2 I shall say more about the implications and relative importance of these different approaches and offer some observations about how they might be knitted together. I will end by discussing the potential applicability of my general approach to answering the “Who?” question to other types of hate speech law.

I. Functional specification

A fourth approach to specifying the proper scope of incitement to hatred laws looks to whatever scope best satisfies the underlying or real functions, purposes or objectives of incitement to hatred laws—functions that can serve to justify in a deep way the very existence of such laws. This is explicitly and directly a matter of connecting the “Who?” question (Who should be protected?) with the “Why?” question (Why hate speech laws?). The nature and content of the relevant underlying or real functions will be contested, of course. Nevertheless, I believe that functional arguments put in favour of incitement to hatred laws can be just as principled as arguments put against such laws.² At any rate, underlying or real functions are to be distinguished from ostensible goals or apparent aims—which

1. Alexander Brown, “The ‘Who?’ Question in the Hate Speech Debate: Part 1: Consistency, Practical, and Formal Approaches” (2016) 29:2 Can JL& Jur 275 [Brown, “Who”].

2. See, e.g., Catharine MacKinnon, *Only Words* (Cambridge: Harvard University Press, 1993) at 83-84; Alexander Brown, *Hate Speech Law: A Philosophical Examination* (New York: Routledge, 2015) at 2 [Brown, “Hate Speech”]. Cf Ivan Hare, “Hate speech” in Peter Cane & Joanne Conaghan, eds, *The New Oxford Companion to Law* (Oxford: Oxford University Press, 2008) 520 at 521.

I have already discussed in relation to what I called practical specification in Part 1—such as to deter acts of incitement to hatred or more simply to punish people who engage in incitement to hatred. Underlying or real functions speak to the fundamental reasons why a society should think it right or justifiable to deter acts of incitement to hatred or to punish people who engage in incitement to hatred in the first place. So what are the real functions of incitement to hatred laws?

A. Symbolic function

One possible function is expressive or symbolic in nature; to send out a message or to represent or signify something important. At face value, the message is simply that the government considers the conduct in question to be unacceptable. But there are also other, deeper messages or symbols that hate speech laws may carry, including, but not limited to, societal disapproval,³ solidarity,⁴ delegitimisation⁵ and even remembrance.⁶ No less importantly, such laws may carry the message or symbolise the idea that people who are subject to hate speech are, in spite of what hate speakers say, “valued members of our polity”.⁷ But if incitement to hatred laws can fulfil valuable expressive or symbolic functions, then perhaps these functions must also dictate to some extent the proper scope of such laws. After all, if the idea that racial minorities, for example, are valued members of the polity is a worthwhile message or symbol to send or carry, worthwhile enough to strengthen the case for laws banning incitement to racial hatred, then surely this could also be a worthwhile message to send out in respect of other groups in the polity, such as groups identified by disability or gender identity, say.⁸ It is important (so the argument would run) for authorities to use the powerfully expressive or symbolic medium of the criminal law to say to these groups, ‘You too are valued members of our polity.’ Of course, none of this precludes governmental authorities from also using other sorts of expressive or symbolic acts to get this message across, including ministerial speeches and pronouncements.

Arguably an official message of equal standing takes on an even greater significance for a particular group when other similar groups do already benefit

3. See, e.g., David F Partlett, “From Red Lion Square to Skokie to the Fatal Shore: Racial Defamation and Freedom of Speech” (1989) 22:3 *Vand J Transnat'l L* 431 at 469; Bhikhu Parekh, “The Rushdie Affair: Research Agenda for Political Philosophy” (1990) 38:4 *Pol Stud* 695 at 705.
4. David Kretzmer, “Freedom of Speech and Racism” (1987) 8:3 *Cardozo L Rev* 445 at 456 [Kretzmer].
5. Anna Elisabetta Galeotti, *Toleration as Recognition* (Cambridge: Cambridge University Press, 2002) at 156.
6. Julie C Suk, “Denying Experience: Holocaust Denial and the Free-Speech Theory of the State” in Michael Herz & Peter Molnar, eds, *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge: Cambridge University Press, 2012) 144 at 154.
7. Mari J Matsuda, “Public Response to Racist Speech: Considering the Victim's Story” (1989) 87:8 *Mich L Rev* 2320 at 2322 [Matsuda].
8. See, e.g., Jonathan Cohen, “More Censorship or Less Discrimination? Sexual Orientation Hate Propaganda in Multiple Perspectives” (2000) 46:1 *McGill LJ* 69 at 81 [Cohen, “More Censorship”]; Kay Goodall, “Challenging Hate Speech: Incitement to Hatred on Grounds of Sexual Orientation in England, Wales and Northern Ireland” (2009) 13:2-3 *Intl JHR* 211 at 223 [Goodall].

from legal protections against hate speech. For, in these circumstances the absence of legal protection could send the message to those who are not protected “that they are second class citizens not entitled to equal protection of the law”.⁹ In the 1990s, for example, some scholars made the case for a new stirring up religious hatred offence in the UK, the function of which would be partly to send a message to Muslims that they were also members of society in equal standing.¹⁰ Later, much the same point was made about the symbolism of a government which declines to provide gays and lesbians with the same legal protections against incitement to hatred afforded to other groups in society.¹¹ More recently, this argument has been applied to disability and transgender identity.¹² In the words of Neil Chakraborti,

laws exist and they send out a powerful symbolic message to those particular communities [who are subject to incitement to hatred on grounds of racial, religious or sexual orientation]. I do not see why the same could not apply when it comes to transphobic hate speech.¹³

Putting this point slightly differently, a failure to legislate against the stirring up of hatred in accordance with the principle of parity (treating like groups alike) might *itself* send out a message of unequal standing or lesser sociolegal status for certain groups in society, depending on other aspects of the context, of course.

Now some people might counter at this stage that functional arguments of the present sort can cut both ways depending on which message is actually being sent out; after all, laws can send out unintended as well as intended messages about equal standing.¹⁴ So, for example, in a society where a great many groups do *not* receive protection from hate speech via incitement to hatred laws, what message is sent out to those few groups who do receive such protection? On one reading, it might say that they are particularly valued members of the polity. But there is an alternative message. Might it say to them that they are too stupid, feckless or cowardly to defend themselves against hate speech? This message is far from edifying and could undermine rather than support the case for having any incitement to hatred laws in that society.

But, then again, maybe it all depends on the group and the context. Surely there is no stigma attached to benefiting from incitement to hatred laws if the relevant laws actually better enable members of the relevant groups to engage in their own counter-speech because they can be confident that the law is there

9. Cohen, “More Censorship”, *supra* note 8 at 75.

10. See, e.g., Tariq Modood, “Muslims, Incitement to Hatred and the Law” in John Horton, ed, *Liberalism, Multiculturalism and Toleration* (Basingstoke: Macmillan, 1993) 139 at 147 [Modood].

11. See, e.g., Marie-France Major, “Sexual-Orientation Hate Propaganda: Time to Regroup” (1996) 11:1 CJLS 221 at 231 [Major].

12. Cf Law Commission, *Hate Crime: Should The Current Offences Be Extended?* (Paper No 348) (London: The Stationery Office, 2014) at para 7.12 [Law Commission].

13. UK, HC, Women and Equalities Committee, “Oral evidence: Transgender Equality Inquiry, HC, 390” (London: The Stationery Office Limited, 2015) at Q46, online: <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/women-and-equalities-committee/transgender-equality/oral/21345.pdf>.

14. Cf Law Commission, *supra* note 12 at paras 7.73-7.89.

to back up or validate what they are saying to hate speakers. Of course, when running this sort of argument it is very important not to make simplistic generalisations about the relative abilities to counter-speak of different groups. To give an example of over-simplification, consider Ian Cram's generalisation that the situation of being less able to counter-speak "may be true for disabled people but not for able-bodied black people, members of religious minorities, gays, lesbians and others."¹⁵ This is misleading for two reasons. On the one hand, arguably Cram goes too far in suggesting a lesser ability to counter-speak on the part of all people with disabilities. By drawing a distinction between 'disabled people' and 'able-bodied black people' Cram seems to conflate being disabled with physical disability. Yet this ignores a potentially salient difference between people with learning disabilities, for whom their disabilities could indeed make it more difficult to engage in counter-speech, and people with physical disabilities, for whom their disabilities might be no greater an impediment to speaking back than being a member of a racial or other minority group. Consider the music of Ian Dury and the Blockheads. On the other hand, Cram does not go far enough in acknowledging the various psychological and emotional hurdles as well as social and material impediments that, for many groups in society, make it difficult for them to actively engage in effective forms of counter-speech, including not merely people with disabilities but also members of racial, religious, sexual orientation, and gender identity minority groups.¹⁶

At any rate, I do not deny that authorities may fail sometimes to accurately predict what message people will take from a law or what will matter to people about a law symbolically. But it does not follow from this lack of certainty that the expressive function of law should be dismissed entirely when determining the proper scope of incitement to hatred laws. Suppose for the sake of argument that what incitement to hatred laws are saying or symbolising to people is the following: 'We the government recognise how challenging it can be for any groups against whom hatred is being stirred up to actively engage in effective forms of counter-speech, so we want to help support this counter-speech by banning incitement to hatred, because we believe that the very existence of such laws will give people the authority as well as the courage they need to speak back.' If this is the real message behind incitement to hatred laws in the UK, then surely it would be much more difficult to justify (on these grounds) a failure to extend the scope of these laws so that they cover not only race, religion, and sexual orientation but also disability and gender identity, for example.

B. Social evils

According to James Weinstein, the expressive or symbolic function of law cannot be the only or main warrant for hate speech laws; only a subsidiary warrant

15. Ian Cram, "Hate Speech and Disabled People: Some Comparative Constitutional Thoughts" in Anna Lawson & Caroline Gooding, eds, *Disability Rights in Europe: From Theory to Practice* (Oxford: Hart, 2005) 65 at 71, n 24.

16. Cf Brown, "Hate Speech", *supra* note 2 at 257-62.

at best.¹⁷ I have discussed this assumption elsewhere and shall not rehearse the arguments here.¹⁸ Instead, I simply point out that the functional approach to scope specification is certainly not exhausted by symbolic or expressive functional arguments. Another highly significant—perhaps even core—function or purpose of hate speech laws is to combat harm. When hate speech takes the form of insults, slurs, derogatory epithets or other abusive comments targeted directly at individuals, especially in face-to-face encounters but also increasingly now via the Internet, the most pressing harms tend to be to the victim’s psychological and physiological health.¹⁹ But when hate speech takes the form of the stirring up of hatred some of the relevant harms are what might be called social evils, that is, evils that characteristically operate at the societal level. On the present way of looking at things, therefore, the proper scope of incitement to hatred laws must reflect, whatever else it reflects, facts about which groups are subject to forms of stirring up hatred that either constitute or contribute to social evils. Of course, there is likely to be some overlap here with practical arguments as well: this is because some functional arguments will assume that incitement to hatred laws are able to combat social evils through the ostensible goal of deterring the relevant forms of hate speech.

What are these social evils? Unsurprisingly, they are many and heterogeneous. One argument appeals to the connection between stirring up hatred and the social evils of discrimination or violence. Defending the government’s introduction of a new offence of stirring up hatred on grounds of sexual orientation in 2007, for example, the then Minister of State for Justice, David Hanson MP, made it clear that the real function or purpose of the offence was not merely expressive but also the deterrence of social evils. “The Government considers that legislation which prohibits the stirring up of hatred will deter such behaviour and send a message that it is unacceptable, leading to homophobic hatred becoming less widespread and in turn reducing the number of incidents of violence, bullying and discrimination.”²⁰ In my view the Minister oversimplified the nature of argument when he declared that “[t]he Government is persuaded that there is a link between the availability of material liable to incite homophobic hatred and levels of homophobic violence.”²¹ It is important not to collapse two steps in the argument: the first is that a high enough incidence of people stirring up hatred against groups can contribute to the creation and maintenance of a climate of hatred toward those groups; the second is that this climate of hatred is characterised in part by a heightened risk of acts of discrimination or violence.²² Be that as it may, what implications does the present argument have for the proper scope of incitement to hatred laws? I take it as read that legislators can be warranted, up to a point, in banning incitement to hatred against a given group if there is minimally

17. See, e.g., James Weinstein, “A Constitutional Roadmap to the Regulation of Campus Hate Speech” (1992) 38:1 Wayne L Rev 163 at 223, 245.

18. See Brown, “Hate Speech”, *supra* note 2 at 241, 248-50.

19. *Ibid* at 49-66.

20. David Hanson, Letter to the Joint Committee on Human Rights, 25 November 2007 at para 6, online: www.publications.parliament.uk/pa/jt200708/jtselect/jtrights/37/3710.htm [Hanson].

21. *Ibid* at para 8.

22. See, e.g., Brown, “Hate Speech”, *supra* note 2 at 66-71.

adequate evidence of the aforementioned links between hate speech and social evils. This seems to be how scholars have viewed the social evils argument for incitement to hatred laws on grounds of race, ethnicity, nationality or religion,²³ sexual orientation,²⁴ and disability.²⁵ I can see no *a priori* reason to rule out the possibility of similar arguments to the effect that new stirring up hatred offences can be warranted up to a point if there is evidence to suggest that people stirring up hatred on grounds of disability or gender identity, say, can contribute to a climate of hatred in which acts of discrimination or violence against the relevant groups are more likely.

A related function of incitement to hatred laws could be to reduce feelings of insecurity suffered by groups who are the subjects of such speech. The relevant working hypothesis is that incitement to hatred can contribute not merely to a climate of hatred but also to a climate of fear.²⁶ In the words of Caroline Flint MP, speaking of the need for a new stirring up religious hatred offence in 2005, “[i]t is widely accepted that individuals in our society are stirring up hatred against particular religious groups” and “[a]t an individual level this can lead to fear and intimidation”.²⁷ As regards to scope, it is often said that members of groups defined by race or religion can be made to feel insecure as a result of the stirring up of hatred against them because of historical patterns in which the stirring up of hatred has been accompanied or proceeded by acts of violence and discrimination.²⁸ It does not require a Herculean leap of imagination to see how other groups in society that have also historically suffered violence and discrimination could also be made to feel insecure by the stirring up of hatred

- 23. See, e.g., Laurence Hauptman, “Group Defamation and the Genocide of American Indians” in Monroe H Freedman and Eric M Freedman, eds, *Group Defamation and Freedom of Speech: The Relationship Between Language and Violence* (Westport: Greenwood Press, 1995) 9 at 11; Alexander Tsesis, *Destructive Messages: How Hate Speech Paves the Way for Harmful Social Movements* (New York: New York University Press, 2002) at 138; L W Sumner, *The Hateful and the Obscene: Studies in the Limits of Free Expression* (Toronto: University of Toronto Press, 2004) at 162-63 [Sumner]; Bhikhu Parekh, “Hate Speech: Is There a Case for Banning?” (2005-6) 12:4 Pub Pol'y Research 213 at 217-18; Glyn Morgan, “Mill's Liberalism, Security, and Group Defamation” in Glen Newey, ed, *Freedom of Expression: Counting the Costs* (Cambridge: Cambridge Scholars Press, 2007) 121 at 136-37 [Morgan]; Alexander Brown, “The Racial and Religious Hatred Act 2006: A Millian Response” (2008) 11:1 Crit Rev Intl Soc & Polt Phil 1 at 13 [Brown, ‘Racial’].
- 24. See, e.g., Major, *supra* note 11 at 227; Cohen, “More Censorship”, *supra* note 8 at 74-75; Martin Bowley, “A Cancer at the Heart of Society” (2000) 150 New LJ 1203 at 1203; Morgan, *supra* note 23 at 136-37; Goodall, *supra* note 8 at 218; Katharine Gelber & Luke McNamara, “The Effects of Civil Hate Speech Laws: Lessons from Australia” (2015) 49:3 Law & Soc'y Rev 631 at 635.
- 25. Mairian Corker, “The UK Disability Discrimination Act: Disabling Language, Justifying Inequitable Social Participation” in Leslie Pickering Francis & Anita Silvers, eds, *Americans with Disabilities: Exploring Implications of the Law for Individuals and Institutions* (New York: Routledge, 2000) 357 at 365 [Corker].
- 26. Brown, “*Hate Speech*”, *supra* note 2 at 71-75.
- 27. Caroline Flint, Letter to the Joint Committee on Human Rights, 3 February 2005 at para 69, online: www.publications.parliament.uk/pa/jt200405/jtselect/jtrights/60/6013.htm.
- 28. See, e.g., Modood, *supra* note 10 at 146; Raymond Chow, “Inciting Hatred or Merely Engaging in Religious Debate? The Need for Religious Vilification Laws” (2005) 30:3 Alt LJ 120 at 120 [Chow]; Brown, “*Hate Speech*”, *supra* note 2 at 71-75. It also seems necessary here that people have good reason to feel insecure, such as if the correlation is non-accidental (*ibid* at 73-74).

against them—for example, gays and lesbians,²⁹ people with disabilities,³⁰ people with transgender identities,³¹ and women.³²

Going beyond the narrow issue of insecurity *qua* fear of discrimination and violence, Jeremy Waldron argues that another important yet overlooked function of hate speech laws (he focuses on group defamation laws) is to curb hate speech which tends to be destructive of a public assurance of civic dignity, resulting in targeted groups feeling insecure as to their equal sociolegal status in society.³³ Here it is relevant to ask not merely ‘Which forms of hate speech are particularly destructive of public assurances?’ but also ‘Which groups are most in need of hate speech laws because of the vulnerability or precariousness of their sense of equal standing in society?’³⁴ Focusing on the latter question, the answer depends in part on the wider context. One could certainly imagine groups of people whose access to rights and liberties is so well established that the mere presence of group defamation, for instance, would do little to reduce their sense of security in their equal sociolegal status.³⁵ By contrast, there may be groups who suffer unequal access to rights and liberties and whose sense of status is more vulnerable to hate speech as a result. Consider the position of gay and lesbian people in Russia who not merely face a lack of protection against incitement to hatred on grounds of sexual orientation, despite the fact that other groups in the country defined by religion, nationality, and race do benefit from such protections,³⁶ but also must endure the existence of laws prohibiting “the promotion of non-traditional sexual relationships” (gay propaganda law).³⁷ Moving on from LGBT issues, what of the position of people with disabilities? This, once again, will depend on the sociolegal position in which they find themselves in a given country, as well as on which types of disabilities are in play. But generalising across contexts and types, it does not seem unreasonable to suppose that if the position of people with disabilities is already precarious—if they already face significant discrimination

- 29. See, e.g., Linda Garnets et al, “Violence and Victimization of Lesbians and Gay Men: Mental Health Consequences” in Gregory M. Herek & Kevin T Berrill, eds, *Hate Crimes: Confronting Violence Against Lesbians and Gay Men* (London: Sage, 1992) 207 at 215-16; Gerald P Mallon, “Sticks and Stones Can Break Your Bones: Verbal Harassment and Physical Violence in the Lives of Gay and Lesbian Youths in Child Welfare Settings” (2001) 13:1-2 J Gay & Lesbian Soc Serv 63 at 71-72; Ellen Faulkner, “Homophobic Hate Propaganda in Canada” (2006) 5:1 J Hate Stud 63 at 81-82 [Faulkner].
- 30. See, e.g., John Stanton-Ife, *Criminalising Conduct with Special Reference to Potential Offences of Stirring Up Hatred Against Disabled or Transgender Persons* (London: Law Commission, 2013) at para 57 [Stanton-Ife].
- 31. *Ibid.*
- 32. See, e.g., Kylie Weston-Scheuber, “Gender and the Prohibition of Hate Speech” (2012) 12:2 QUT L & Just J 132 at 141-43 [Weston-Scheuber].
- 33. Jeremy Waldron, “Dignity and Defamation: The Visibility of Hate” (2010) 123:7 Harv Rev 1596 at 1623, 1630 [Waldron].
- 34. Brown, “Hate Speech”, *supra* note 2 at 148-56.
- 35. *Ibid* at 150-152.
- 36. *The Criminal Code of the Russian Federation*, 1996/63, art 282.1.
- 37. Russian Federation, 24 December 2010, *Federal Law No. 436-FZ of 2010 on Protection of Children from Information Harmful to their Health and Development*, as amended by *Federal Law No.135-FZ of 2013, on Amendments to Article 5 of the Federal Law “On the Protection of Children from Information Harmful to their Health and Development” and Miscellaneous Legislative Acts of the Russian Federation for the Purpose of Protecting Children from Information that Promotes the Rejection of Traditional Family Values*.

in the workplace and discrimination in the provision of public services, such as housing, access to public spaces, and so on, and if anti-discrimination laws, such as they exist, lack teeth or are not properly enforced—then it is quite possible that, as Stanton-Ife puts it, “[e]ven a small or incipient phenomenon of persons intending to stir up hatred against these groups would suffice for the applicability of Waldron’s argument.”³⁸

An additional function of incitement to hatred laws might be to maintain public order in the traditional sense of keeping the peace. In England and Wales this function is indicated by the fact that the stirring up hatred offences have been classified as public order offences via the *Public Order Act 1986*.³⁹ If maintaining public order is one of the underlying or real functions of incitement to hatred laws, then it might be tempting to try to draw distinctions between different protected characteristics.⁴⁰ It has been suggested by some people, for example, that stirring up hatred on grounds of sexual orientation, disability or gender identity is not as significant a threat to public order as stirring up racial or religious hatred.⁴¹ As a brute statistical fact about the percentage of public disturbances that historically have had something to do with racial or religious identity as compared to other markers of identity this might turn out to be true. But, then again, this argument could also point in unexpected directions. In some countries and at certain times, the logic of safeguarding public order might warrant incitement to hatred laws that covered social class or even profession as protected characteristics. Such laws were once part of the Italian Criminal Code,⁴² and it is not beyond the realms of possibility that a future Italian government could reintroduce them, perhaps in response to events like the international ‘day of rage’ in 2011 during which peaceful protesting against bankers and politicians descended into violence between protestors and police, rioting, vandalism, and arson.

- 38. Stanton-Ife, *supra* note 30 at para 63. Once again, it seems necessary here that people have good reason to feel insecure in their sociolegal status (*ibid* at paras 70-71).
- 39. (UK), c 64, parts 3, 3A [*Public Order Act 1986*]. Note, the stirring up hatred offences can be traced back to the criminal offence of seditious libel in English common law. But it is important to recognise that the criminal offence of seditious libel evolved into two distinct clusters of laws. First, laws that ban the stirring up of hatred. See, e.g., Brown, “*Hate Speech*”, *supra* note 2 at 26-28. Second, laws that prohibit public speech or other expressive conduct concerning or directed at members of groups or classes of persons identified by protected characteristics that constitutes a threat to public order, in the sense that it causes or is likely to cause a public mischief, breach of the peace, or danger to public safety. See, e.g., *ibid* at 28-29. Pace Lord Justice Lawton in *R v Edwards* (1983) 5 Cr App R (S) 145 at 148, I would argue that both clusters have an equally strong claim to being the modern day equivalent of the criminal offence of seditious libel.
- 40. Distinctions might also be drawn between different kinds of speech. For example, it could conceivably be argued that criticism or insult of religious prophets or religious beliefs (where this does not amount to stirring up religious hatred) poses an even greater threat to public order than stirring up religious hatred because religionists may feel that they have a religious duty to react strongly to such criticisms or insults. Cf British Humanist Association, Final (Third) Submission to the Select Committee on Religious Offences: Legislation on Incitement of Religious Hatred, 4 December 2002 at para 6, online: <https://humanism.org.uk/wp-content/uploads/BHA-Submission-on-Religious-Offences-3.pdf>.
- 41. See, e.g., Simon Bronitt & James Stellios, “Sedition, Security and Human Rights: ‘Unbalanced’ Law Reform in the ‘War on Terror’” (2006) 30:3 Melbourne UL Rev 923 at 947.
- 42. David Riesman, “Democracy and Defamation: Control of Group Libel” (1942) 42:5 Colum L Rev 727 at 744-45.

Then again, perhaps what matters here is not statistics on public disturbances but deeper sociological explanations about why hate speech against certain characteristics might be especially dangerous. Maybe the thought is that race and religion are special because they are closely related to interpersonal attraction, family-making and group formation—all of which can be focal points for powerful psychological drives to stick together and protect one’s own, including, if necessary, through the use of violence. Yet surely for many people their sexual orientation, gender identity, or disability is also closely related to interpersonal attraction, family-making and group formation. At any rate, it would be naive, I think, to simply assume that the threat-levels associated with incitement to hatred on grounds of sexual orientation, disability or gender identity could never be high enough to justify the creation of new stirring up hatred offences. Consider a small group of people distributing leaflets or holding up banners or placards inciting hatred against transgender people directly outside a building in which a transgender beauty pageant is taking place. Could we confidently say that their actions could never pose a threat to public order? Here the threat of violence might derive not from people who agree with the transphobic hate speech (sympathisers) but from those who object to it (hecklers).

Of course, those who defend incitement to hatred laws must also confront the heckler’s veto objection: namely, that facts about how an audience is likely to respond to speech, such as by disturbing the peace, should not be used as a basis for restricting speech since it would in effect give audiences the power to censor speech by threat of disturbing the peace.⁴³ I do not intend to try to rebut that argument here. Instead, I simply want to make the point that *if* threats to public order from predictable audience reactions *are* a sound basis for, or can warrant, incitement to hatred laws, despite the heckler’s veto objection, then this basis should hold for any characteristics concerning which such audience reactions can be predicted. It would be an inconsistent argument that said a threat of public disturbance relating to the stirring up of racial or religious hatred can be a basis for prohibition but a similar threat of similar public disturbance relating to the stirring up of hatred on grounds of sexual orientation, disability or gender identity cannot.

Yet another, admittedly very broad, family of social evils that a society might choose to combat through hate speech laws can be described in terms of the egalitarian effects of such speech. In some contexts, hate speech can (i) reinforce societal and institutional prejudices, (ii) deprive groups of important rights and powers, such as by inhibiting normal functioning within, or even forcing a complete withdrawal from, educational settings, the workplace, leisure and exercise facilities, civil society, and the political sphere or public forums broadly construed, (iii) enact or perpetuate forms of subordination, oppression, and political powerlessness, or (iv) support and maintain a background of social norms that in turn can make it more difficult for groups to engage in effective counter-speech.

43. See Harry Kalven, Jr, *The Negro and the First Amendment* (Chicago: University of Chicago Press, 1965); David AJ Richards, *Tolerance and the Constitution* (New York: Oxford University Press, 1986); David AJ Richards, *Free Speech and the Politics of Identity* (Oxford: Oxford University Press, 1999) [Richards, “Free Speech”].

Writers have linked such effects not merely with hate speech concerning race, ethnicity, nationality, and religion⁴⁴ but also with hate speech concerning sex or gender identity, including female identity⁴⁵ and transgender identity,⁴⁶ sexual orientation, mainly being gay or lesbian,⁴⁷ disability,⁴⁸ and physical appearance, most notably obesity.⁴⁹ Of course, not all the writers who recognise these inegalitarian effects believe that the right response to them, all things considered, is to enact criminal incitement to hatred laws. Clearly some do not.⁵⁰ But for those writers who do think that there is at least a *pro tanto* case to be made for such laws, they have had no conceptual or empirical difficulty in extending their arguments to a broad range of characteristics.

However, at this point it might be suggested that hate speech is likely to produce differential effects, in terms of causing or not causing forms of social inequality and oppression, depending on which protected characteristics are at stake. Take (iv) which has to do with how hate speech can help to support and maintain social norms that can block counter-speech, such as a norm that says victims of hate speech ought to turn the other cheek. Interestingly, however, in her study of the motives of hate speakers and of the response strategies adopted by people who are the subjects of hate speech in the US, Laura Leets set out to test, amongst other things, the hypothesis that “Jews will report more assertive response strategies than homosexuals” potentially because “Jews have a more secure place in society than homosexuals” and because “increased assertiveness on [the part of homosexuals] may have a more perilous effect (e.g., antigay violence) than for their Jewish counterparts”.⁵¹ If proven to be true, this hypothesis

- 44. See, e.g., Kent Greenawalt, *Speech, Crime, and the Uses of Language* (New York: Oxford University Press, 1989) at 148 [Greenawalt, “Speech”]; Kent Greenawalt, *Fighting Words: Individuals, Communities and Liberties of Speech* (Princeton: Princeton University Press, 1995) at 59-60 [Greenawalt, “Fighting Words”]; Chow, *supra* note 28 at 120-21; Brown, “Hate Speech”, *supra* note 2 at 80-86, 194-201, 257-62.
- 45. See, e.g., Greenawalt, “Speech”, *supra* note 44 at 299; Greenawalt, “Fighting Words”, *supra* note 44 at 59-60; Donna L Lillian, “A Thorn By Any Other Name: Sexist Discourse as Hate Speech” (2007) 18:6 Discourse Soc 719 at 736.
- 46. See, e.g., Martha T Zingo, *Sex/Gender Outsiders, Hate Speech and Freedom of Expression: Can They Say That About Me?* (Westport: Praeger, 1998) at 165-67 [Zingo].
- 47. See, e.g., Greenawalt, “Speech”, *supra* note 44 at 299; Greenawalt, “Fighting Words”, *supra* note 44 at 59-60; Zingo, *supra* note 46 at 165; Cohen, “More Censorship”, *supra* note 8 at 94; Faulkner, *supra* note 29 at 81-82; Eric Heinze, “Cumulative Jurisprudence and Hate Speech: Sexual Orientation and Analogies to Disability, Age, and Obesity” in Ivan Hare & James Weinstein, eds, *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009) 265 at 279 [Heinze, “Cumulative Jurisprudence”].
- 48. See, e.g., Corker, *supra* note 25 at 366; Heinze, “Cumulative Jurisprudence”, *supra* note 47 at 276; Timothy Shriver, “The Bigotry Behind the Word ‘Retard’”, *The Washington Post* (15 February 2010), online: www.washingtonpost.com/wp-dyn/content/article/2010/02/14/AR2010021402893.html; Andreas Dimopoulos, “Balancing Disability Protection Against Freedom of Speech: Should an Offence of Incitement to Disability Hatred be Introduced?” (2015) Pub L 79 at 92-93.
- 49. See, e.g., UK, All Party Parliamentary Group on Body Image, *Reflections on Body Image* (London: HMSO, 2012) at 61, online: www.ncb.org.uk/media/861233/appg_body_image_final.pdf.
- 50. Zingo, *supra* note 46; Heinze, “Cumulative Jurisprudence”, *supra* note 47; Eric Heinze, *Hate Speech and Democratic Citizenship* (Oxford: Oxford University Press, 2016).
- 51. Laura Leets, “Experiencing Hate Speech: Perceptions and Responses to Anti-Semitism and Antigay Speech” (2002) 58:2 J Soc Issues 341 at 345-46 [Leets].

might have suggested that for Jews, unlike gays, hate speech laws (potentially including incitement to hatred laws) would not be warranted since Jews, unlike gays, could adopt an alternative remedy of assertively talking back to hate speakers. However, Leets discovered that this hypothesis was not supported by the results of her research. In fact, the data “indicated no significant difference across the two groups [in the extent of assertive response strategies]”.⁵²

C. Interests-balancing

Another functional strategy might be to specify the scope of incitement to hatred laws in whichever way achieves the optimum balance of interests, including interests in freedom of expression, where optimality might be defined in terms of whichever scope maximises the net satisfaction of interests.⁵³ How would this approach play out in practice? What distinctions would it suggest, if any, between incitement to hatred laws which protect different characteristics? One possibility is that authorities should distinguish between incitement to *religious* hatred laws and incitement to *racial* hatred laws to reflect differential risk factors that these laws will be overbroad (ban valuable speech) or have a chilling effect (deter valuable speech).⁵⁴ In other words, if laws banning incitement to religious hatred pose a much greater threat to interests in freedom of expression than do laws banning incitement to racial hatred, then, under an interests-balancing approach to the “Who?” question, this might well tip the scales against the former but in favour of the latter.

In fact, precisely this argument came to prominence in England and Wales around the time of the *Racial and Religious Hatred Bill*.⁵⁵ Several commentators argued that those people tasked with enforcing the proposed new legislation banning the stirring up of religious hatred would have an especially hard time distinguishing between speech that stirs up hatred against religious *believers* and speech that merely stirs up hatred against religious *beliefs* (such as through criticism, ridicule and satire).⁵⁶ The legislation will unavoidably (so the objection

52. *Ibid* at 353.

53. Cf Sumner, *supra* note 23 at 60-63; Evan Simpson, “Responsibilities for Hateful Speech” (2006) 12:2 Leg Theory 157 at 158; Jonathan Gilmore, “Expression as Realization: Speakers’ Interests in Freedom of Speech” (2011) 30:5 Law & Phil 517 at 539; Frederick Schauer, “Social Epistemology, Holocaust Denial, and the Post-Millian Calculus” in Michael Herz & Peter Molnar, eds, *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge: Cambridge University Press, 2012) 129 at 138; Brown, “Hate Speech”, *supra* note 2 at 222-34.

54. See, e.g., Susannah C Vance, “The Permissibility of Incitement to Religious Hatred Offenses Under European Convention Principles” (2004) 14:1 Transnat'l L & Contemp Probs 201 at 205 [Vance]; Ivan Hare, “Crosses, Crescents and Sacred Cows: Criminalising Incitement to Religious Hatred” (2006) Pub L 521 at 537-538 [Hare, “Crosses”].

55. Bill 31, 2005-2006 sess, 2005 [*Racial and Religious Hatred Bill*].

56. See, e.g., Rowan Atkinson, “The Opposition’s Case” in Lisa Appignanesi, ed, *Free Expression is No Offence* (London: Penguin, 2005) 59 at 60; Gareth Crossman, “Religious Hate: A Criminal Offence? Liberty Says No” (2005) Leg Action 9 at 9; Philip Pullman, “Against ‘Identity’” in Lisa Appignanesi, ed, *Free Expression is No Offence* (London: Penguin Books, 2005) 105 at 110. Cf Eric Barendt, “Religious Hatred Laws: Protecting Groups or Belief?” (2011) 17:1 Res Publica 41.

went) either ban or chill the latter along with the former, and thereby ride rough shod over weighty speaker and audience interests in freedom of expression on religious affairs. Of course, originally the *Racial and Religious Hatred Bill*⁵⁷ did not contain the absurd rule of thumb that any attempt to stir up hatred against a set of religious beliefs shall be taken as *prima facie* evidence of an intent to stir up hatred against religious believers. Even so, the worry seemed to be that the police, public prosecutors, magistrates, judges and juries could end up unwittingly following this rule of thumb. And so the worry was as much about legal practitioners as about the law: about *their* ability to make sensible decisions, based on the facts of the case, as to whether a suspect or defendant had stirred up hatred against religious believers as opposed to against religious beliefs.

However, it should not be forgotten that legal professionals already had to distinguish between speech that stirs up hatred against people due to their race or ethnicity and speech that merely stirs up hatred against aspects of racial or ethnic identity and culture itself. To see this point consider the following example. Suppose for the sake of argument that legal professionals would have a hard time (although I can find scant evidence to suggest that they would) distinguishing between ‘You think you can live with Muslims, think again, they are vile, backward, and sadistic people who deserve only our hatred, and when this country is finally united in its hatred of them they had better watch out!’ and ‘Unlike Christianity, Islam doesn’t teach that Jesus died on the cross for the sins of humankind and doesn’t teach that our eternal salvation lies in having faith in Jesus Christ; instead Islam teaches that salvation is earned through faithful obedience to the teachings of Allah, as revealed through the Prophet Muhammad, and in my opinion this makes Islamic beliefs more dangerous than Christian beliefs.’ Would legal professionals have any easier or harder time distinguishing between ‘I hate these golliwogs who come over here and take our jobs and if you cared about this country you would hate them as well, so trust me when I say that if we all start listening to our true feelings about these people they had all better be careful!’ and ‘White British ethnicity includes a cultural heritage of cricket and cucumber sandwiches, whereas Afro-Caribbean British ethnicity includes a cultural heritage of chaotic street carnivals and recreational marijuana use, and in my opinion this makes Afro-Caribbean British ethnicity more pernicious than White British ethnicity’?

Notwithstanding this last point, the aforementioned worry about how legal professionals might come to enforce laws banning incitement to religious hatred was subsequently addressed head-on by members of the House of Lords (notably Lords Hunt, Lester, Carey, and Plant). In October 2005 the Lords succeeded in amending the legislation so as to better protect free speech interests. Thus, the *Public Order Act 1986* now contains s 29J.

Protection of freedom of expression: Nothing in this Part shall be read or given effect in a way which prohibits or restricts discussion, criticism or expressions of antipathy, dislike, ridicule, insult or abuse of particular religions or the beliefs or practices of their adherents, or of any other belief system or the beliefs or practices

57. *Supra* note 55.

of its adherents, or proselytising or urging adherents of a different religion or belief system to cease practising their religion or belief system.⁵⁸

What this seems to show is that laws banning incitement to religious hatred may require very careful drafting and a relatively narrow specification in order to minimise any negative impact on freedom of expression.⁵⁹ What it does not show, however, is that laws banning incitement to religious hatred are invariably and irredeemably *more* hostile to free speech interests than other similar laws.⁶⁰ After all, a similar need for careful drafting and narrow specification was present when the government introduced a new offence of stirring up hatred on grounds of sexual orientation and no doubt the same need would also be present if the government sought to create yet further stirring up hatred offences to cover disability and gender identity.⁶¹

To expand on these last observations, when the government introduced the *Criminal Justice and Immigration Bill*⁶² concerns were raised that legal professionals would find it hard to distinguish between speech that constitutes stirring up hatred on grounds of sexual orientation and speech that is merely religiously motivated criticism of homosexual conduct. Yet Lord Waddington successfully put forward an amendment that was designed to limit the perceived threat to relevant free speech interests, in a parallel fashion to s 29J of the *Public Order Act 1986*. Consequently, the *Public Order Act 1986* now also contains s 29JA.

Protection of freedom of expression (sexual orientation): (1) In this Part, for the avoidance of doubt, the discussion or criticism of sexual conduct or practices or the urging of persons to refrain from or modify such conduct or practices shall not be taken of itself to be threatening or intended to stir up hatred. (2) In this Part, for the avoidance of doubt, any discussion or criticism of marriage which concerns the sex of the parties to marriage shall not be taken of itself to be threatening or intended to stir up hatred.⁶³

Thus far the courts have not found it impossible in practice to distinguish between speech that stirs up hatred against homosexuals and speech that merely expresses religionists’ interpretations of religious texts on the subject of homosexual conduct.⁶⁴ Intriguingly, the same arguments are now being played out *vis-à-vis* the

58. *Supra* note 39, as amended by the *Racial and Religious Hatred Act 2006* (UK), c 1 [*Racial and Religious Hatred Act 2006*].

59. Cf Ian Leigh, “Homophobic Speech, Equality Denial, and Religious Expression” in Ivan Hare & James Weinstein, eds, *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009) 375 at 387.

60. Cf Hare, “Crosses”, *supra* note 54 at 537-38.

61. See, e.g., Stanton-Ife, *supra* note 30 at paras 66, 84-85.

62. Bill 16-I, 2007-2008 sess, 2008.

63. *Supra* note 39, as amended by the *Criminal Justice and Immigration Act 2008* (UK), c 4 [*Criminal Justice and Immigration Act 2008*] and the *Marriage (Same Sex Couples) Act 2013* (UK), c 30 [*Marriage (Same Sex Couples) Act 2013*].

64. Consider the English case *R v Ali, Javed, and Ahmed* (2012) No. T20110109, Derby Cr. Ct., 10 February. In July 2010 three devout but also socially conservative members of the Muslim faith distributed leaflets on the streets of Derby titled ‘Turn or Burn’, ‘GAY—God Abhors You’, ‘Death Penalty?’ as a protest to the Gay Pride Festival taking place that day. They became the first people to be successfully prosecuted for offences relating to stirring up hatred on grounds of sexual orientation in England and Wales. In his sentencing remarks Judge Burgess made reference to clause 29JA (that denouncing homosexual practices as immoral shall not

possibility of a third major extension of the stirring up hatred offences in England and Wales to cover disability and gender identity. Some of the individuals and organisations who participated in the Law Commission's consultation exercise on this issue argued that "such new offences would stifle legitimate criticism of gender reassignment surgery".⁶⁵ However, this only invites the following obvious response: why not simply introduce bespoke freedom of expression clauses for disability and gender identity? Let us suppose for the sake of argument that legal professionals could, without too much difficulty, appeal to the freedom of expression clause s 29JA, along with other elements of the offence of stirring up hatred on grounds of sexual orientation, in order to distinguish between 'We must all hate gays because that is all they deserve as much for the HIV epidemic they brought to the world as for the rampant paedophilia they are responsible for, and when we finally come together in our hatred of gays they had better watch out!' and 'I urge you to stop engaging in homosexual conduct', and in order to distinguish between 'If you want to be at peace with God you must all hate gays as much as God hates gays, and when we finally come together in our hatred of gays they had better watch out!' and 'My religion tells me that homosexual acts are sinful in the eyes of God.' Surely legal professionals could also appeal to a bespoke freedom of expression clause, along with other elements of a new offence of stirring up hatred on grounds of gender identity, say, in order to distinguish between 'I tell you that we as real women have no option but to hate and despise so-called trans women because they are at the vanguard of ever-more extreme forms of male dominance over women, and I can also promise you that when we as real women do unite in our hatred of these imposters they had better watch out!' and 'Female sex/gender is a complex identity that depends on innumerable distinctive childhood and adult experiences that are themselves shaped by a mixture of biology and cultural context in myriad ways that could never be achieved or simulated by sex reassignment surgery and hormone therapy alone.'

Of course, some people may find any such clauses unsatisfactory as mitigation strategies in response to the threat to free speech interests posed by incitement to hatred laws. But then the question becomes whether there is anything unique or special about disability and gender identity as far as these mitigation strategies are concerned. Is there any basis on which to think that inserting bespoke freedom of expression clauses into the relevant legislation concerning stirring up of hatred on grounds of disability and on grounds of gender identity will be inherently less effective in mitigating the threat to valuable speech than the similar clauses employed in the existing legislation concerning stirring up hatred on grounds of religion and on grounds of sexual orientation? If there is such a basis, I cannot see what it is. So if there are sound bases for rejecting the freedom of expression clause strategy for mitigating the threat to free speech interests, then surely these grounds would apply equally to a range of protected

be taken of itself to be threatening or intended to stir up hatred) but nevertheless supported the jury's decision that in this particular case the wording of the leaflets *did* amount to the use of threatening words or behavior with the intention of stirring up hatred. Transcript obtained directly from Judge Burgess.

65. Law Commission, *supra* note 12 at para 7.41.

characteristics, perhaps all characteristics, and not just to disability and gender identity uniquely.⁶⁶

A related worry might be that it is especially difficult in the case of religion as opposed to race for legal professionals to distinguish between speech that is intended to *stir up* hatred against a group or class of persons and speech that is simply intended to *express* hatred for that group or class of persons. But, once again, it is hard to see the rational basis for this worry. Compare a person who stands on Oxford Street with a megaphone proclaiming ‘Listen to me, you must understand that all Muslims are wife-beaters and jihadists and as such deserve only our hatred and hostility, and trust me when I say that they should tremble at what is in store for them when we finally open our eyes to who they really are’ with a person who stands on Regent Street with a megaphone declaring ‘Deep down you know that all niggers are layabouts and thieves and so please don’t be afraid to hate these people because hating them will be the first step along a patriotic path to sending them back to the jungle with their tails behind them, and at the point of a gun.’ It is a matter for the interpretation of courts, based on the contexts of utterance and the states of mind of the speakers, whether either of these utterances are deliberate attempts to stir up hatred or merely attempts to express hatred. But it strikes me that the first case is no lesser a clear-cut example of stirring up hatred than the second case despite the fact that the first case concerns the religion of the targeted group and the second case concerns the race of the targeted group. I do not deny that these sorts of cases may pose significant challenges for the police, prosecutors, magistrates, judges and juries. Rather, my point is that there does not seem to be any difference in propositional content that would make stirring up hatred inherently more difficult to distinguish from expressing hatred in the case of religion than in the case of race.

But perhaps I am missing the point. Maybe the real concern is not that it is harder to distinguish between the speech act of stirring up hatred against other people and alternative types of speech act in the case of religion than in the case of race but instead that the value of speech which is used to stir up hatred tends to be higher in the case of religion than in the case of race and this alters the balance of interests. How so? One thought might be that the words people use to stir up religious hatred fall into the category of political speech, whereas words used to stir up racial hatred do not. And since political speech should be protected due to democratic interests we all share (so the argument goes), the case for incitement to religious hatred laws is harder to make than for incitement to racial hatred laws.⁶⁷ However, it is difficult to sustain this distinction as soon as one reflects on the fact that very often people also use highly politicised racist hate propaganda in order to stir up racial hatred.⁶⁸ Then again, a second, even more general distinction is that when it comes to stirring up racial hatred nothing is ever said

66. Cf *ibid* at para 7.43.

67. Cf Vance, *supra* note 54 at 205.

68. See, e.g., *R v Birdwood* (1995) No. 94/2421/X2, EWCA Crim, 11 April (involving the dissemination of an anti-Semitic pamphlet to all MPs, MEPs and all those on the Church of England’s yearbook list).

that deserves to be protected—because, for example, the words or sentences used to stir up racial hatred always either express malicious falsehoods on matters of fact or amount to nothing more than mindless outpourings of emotion that fail to connote any actual ideas or opinions—whereas what passes for stirring up of religious hatred is invariably worth protecting—because, for example, the relevant words or sentences never express malicious falsehoods on matters of fact and are always vessels for actual ideas and opinions, such as ideas about how certain people's commitment to their religious doctrines causes them to behave in savage or uncivilised ways.⁶⁹ But, once again, the distinction seems unfounded. On the one hand, surely the stirring up of racial hatred can, and often does, take the form of words or sentences that connote actual ideas or opinions, including ideas about the alleged moral inferiority of certain races.⁷⁰ On the other hand, it is surely not difficult to imagine instances of stirring up religious hatred involving words or sentences which are either malicious falsehoods on matters of fact ('Each and every Muslim on this street is directly involved in plotting terrorist attacks') or mindless outpourings that do not connote any actual ideas or opinions but instead express non-cognitive attitudes or feelings such as disapproval, contempt or hatred ('Damn Muslims!').

So far I have talked about the balancing of interests with a special emphasis on the interests of speakers and audiences. Another potential way of approaching this sort of balancing is to place extra weight on the interests of those who are the objects of the stirring up of hatred. This is about weighing up the extent to which members of groups identified by given characteristics have a greater interest in legal protections against people stirring up hatred against them or a greater interest in enjoying unqualified freedom of expression themselves, including the freedom to engage in speech that potentially could be construed by the police and the courts as stirring up hatred against others. This is likely to involve highly contextual judgements, depending on the group, the society and the moment in history. In the case of racial minorities in the US, for example, several scholars have highlighted the way in which historically oppressed citizens have utilised their right to freedom of expression enshrined in the First Amendment to campaign for civil rights legislation that prohibits religious, racial, ethnic, and other forms of discrimination.⁷¹ On one interpretation of these events, for these groups to now claim protection through hate speech

69. See, e.g., Vance, *supra* note 54 at 244; Randall Hansen, "The Danish Cartoon Controversy: A Defence of Liberal Freedom" (2006) 44:5 *Int'l Migration* 7 at 12; Hare, "Crosses", *supra* note 54 at 534.

70. See, e.g., *R v Sheppard and Whittle* (2009) No. T20080094, Leeds Cr. Ct., 10 July (involving the possession, publication and distribution in print and through a website of material expressing the moral inferiority of certain races).

71. See, e.g., Samuel Walker, *In Defense of American Liberties: A History of the ACLU* (New York: Oxford University Press, 1990); Samuel Walker, *Hate Speech: The History of an American Controversy* (Lincoln: University of Nebraska Press, 1994); William B Rubenstein, "Since When is the Fourteenth Amendment Our Route to Equality? Some Reflections on the Construction of the 'Hate-Speech' Debate from a Lesbian/Gay Perspective" in Henry Louis Gates Jr et al, eds, *Speaking of Race, Speaking of Sex: Hate Speech, Civil Rights, and Civil Liberties* (New York: New York University Press, 1994) 280 [Rubenstein, "Since"]; Richards, "Free Speech", *supra* note 43.

laws would be, first, a gross betrayal of the very constitutional essential that enabled them to achieve the advances they have made and, second, ill-advised given that they may continue to need protection under the First Amendment in order to campaign vigorously for other issues in the future and to prevent any retrenchment of hard won advances.⁷² As Martha Zingo puts it, “the harms arising from hate speech may involve less risk for lesbians, gay men, bisexuals, and transgenderists than the potentially greater threat (in both kind and duration) posed by censoring speech.”⁷³

It is certainly true that any group claiming protection through hate speech laws also risks having those same or similar laws turned against their own speech. Nevertheless, I think it right to acknowledge that there is a genuine dilemma or quandary for groups who find themselves the objects of hate speech. From the perspective of these groups there are benefits as well as risks associated with hate speech laws. Consider sexual orientation. “While restrictions on hate propaganda might be manipulated in discriminatory ways, such restrictions might also embolden sexual minorities to express their sexuality publicly.”⁷⁴ Moreover, there are risks associated with constitutional rules disallowing hate speech laws as well as risks associated with hate speech laws themselves. Thus, “just as rules of law allowing the repression of speech are used against lesbians, so too are rules of law forbidding the repression of speech.”⁷⁵ Therefore, “[t]hat lesbians and others fighting against the repression of lesbian speech often rely on the rule of law’s enshrinement of free speech should not necessarily be determinative of issues relating to hate crimes involving speech.”⁷⁶ Speaking now of the UK specifically, I do not believe that it has to be the case, nor should it be the case, that being the object of speech that stirs up hatred is a price that people must pay in order to enjoy the right to freedom of expression. On the contrary, it is quite possible that narrowly specified and sensibly enforced stirring up hatred offences could protect people from gratuitous hate speech without at the same time gratuitously infringing their own right to freedom of expression. But even if people do face a genuine quandary, this quandary is not, at least in the UK at the present time, substantially different for people identified by their disability or gender identity as it is for people identified by their race, religion or sexual orientation. In other words, if there is a risk that people will fall foul of the very stirring up offences that were designed to protect them, at first glance this risk does not seem any greater or lesser for people with disabilities or people with transgender identities than it does for Jews, Muslims, gays and lesbians.

Reflecting on all this, what should we say about the functional approach in general? Two problems are worth highlighting. The first problem is that a given hate speech law might secure multiple functions and in turn those different

72. See, e.g., Rubenstein, “Since”, *supra* note 71 at 21-22; Nadine Strossen, “Incitement to Hatred: Should There Be a Limit?” (2001) 25:2 S III ULJ 243 at 262-66.

73. Zingo, *supra* note 46 at 178.

74. Cohen, “More Censorship”, *supra* note 8 at 92.

75. Ruthann Robson, *Lesbian (Out)law: Survival Under the Rule of Law* (Ithaca: Firebrand Books, 1992) at 154 [Robson].

76. *Ibid.*

functions might support drawing the scope of that law differently. In other words, there could be indeterminacy of judgment when using this approach by itself. This problem mainly stems from the fact that the functional approach is itself an umbrella strategy covering a variety of different functional arguments. To illustrate the problem, let us suppose for the sake of argument that the aim of optimally balancing interests supports the enactment of laws banning incitement to racial hatred but not laws banning incitement to religious hatred. (I am sceptical about this, however let us assume it is true in spite of my scepticism.) What if other functional arguments suggest a different strategy? It could be that having laws on the books banning both incitement to racial hatred and incitement to religious hatred helps to send out a message that both racial and religious groups, both Jews and Muslims, for instance, are valued members of society. Or consider once more the functional arguments which say that incitement to religious hatred laws no less than incitement to racial hatred laws can be important tools in combating speech that contributes to climates of hatred and fear—climates in which there is an increased risk of acts of discrimination or violence and a heightened sense of insecurity. The upshot is that even if some functional arguments would differentiate between the stirring up of racial hatred and the stirring up of religious hatred, other functional arguments might not. In other words, whereas some functional arguments may permit exceptions to the principle of parity (treating like groups alike), others might point in the direction of upholding the principle of parity.

A second problem—or set of problems—stems from the fact that the internal logic of a purely functional approach to the issue of scope seems to imply moving beyond broadly defined characteristics such as race, religion, sexual orientation, disability, gender identity, and so on, in order to differentiate between relevant sub-groups. Not race but particular races; not religion but particular religions; not sexual orientation but particular orientations; not disability but particular disabilities; not gender identity but particular gender identities; and so on. For example, one possible reason for limiting the scope of incitement to hatred laws to particular sub-groups rather than the broader characteristics themselves has to do with the two step argument concerning the climate of hatred touched upon earlier. The first step links stirring up hatred to the climate of hatred. One can certainly imagine, for example, a scenario in which, due to the high incidence of speech that stirs up hatred against Muslims or due to the pre-existing conditions that are conducive to such speech producing a climate of hatred against Muslims, marginal instances of stirring up hatred against this particular sub-group make a larger contribution to, or do more to accelerate the growth of, the climate of hatred than general instances of stirring up hatred on grounds of the broader characteristic, religion. The fact that there is this greater contribution or tipping point could conceivably support limiting the scope of the relevant legislation to just the sub-group. The second step in the argument links the climate of hatred to an increased risk of acts of discrimination or violence. Once again, it could turn out to be the case that the climate of hatred which surrounds the particular sub-group, Muslims, comprises an especially

heightened risk of acts of discrimination or violence as compared to the general climate of hatred, if, indeed, there is one, surrounding the broad characteristic, religion. Again, this heightened risk may warrant a legislative focus on the sub-group. The problem is, however, that the splintering of this body of law may attract a series of objections on the part of overlooked sub-groups that do not receive similar protections: namely, that hate speech laws violate the principle of parity. Lack of parity might also threaten the public good of political stability. It is surely feasible that a government regime that offers protection against hate speech to some religious sub-groups but not to others may be a little bit less stable than one that ensures protection for all religious sub-groups. These same problems are likely to arise when there is differential treatment of sub-groups under the broader categories of race, ethnicity, sexual orientation, disability, gender identity, and so on. Lack of parity may also breed or exacerbate broader social tensions between sub-groups. Specifically, it might provide people with yet another source of motivation for engaging in hate speech, and yet another source of inspiration for innovations in the shape and content of that hate speech going forward. In this way, choosing to protect a particular sub-group against incitement to hatred because they are especially susceptible to the relevant social evils might turn out to be a type of self-fulfilling prophecy, and not in a good way.

Related to this problem are various feasibility issues surrounding the introduction of new hate speech laws. Recall from Part 1 the list of hate speech laws grouped together based on the characteristics being protected. It is noticeable that some laws protect HIV/AIDS status (e.g., Australian Capital Territory (Australia),⁷⁷ New South Wales (Australia)⁷⁸) rather than medical status in general, including suffering from any series disease (e.g., Facebook⁷⁹). Many laws specify disability as a protected characteristic rather than the more general category of human capability or functioning which includes ability or disability. As a result it remains permissible for people with disabilities to stir up hatred against other people on grounds of their being TABs or temporarily able-bodied. Some university anti-harassment policies (or campus speech codes) protect veteran status but do not seem to cover discriminatory harassment against people who are pacifists. In the case of sex or gender identity, some laws do not specify which sub-groups are protected meaning that, in theory, all sub-groups are protected including man, woman, male, female, cisgender, transgender, cissexual, transsexual, third gender, bigender, pangender, agender, intersex, third sex (e.g., Tasmania (Australia)⁸⁰), whereas others specify only transgender (e.g., New South Wales (Australia)⁸¹). Some laws appear to protect various forms of sexual orientation including heterosexual, homosexual, lesbian, bisexual, and

77. *Discrimination Act 1991* (ACT), ss 65, 66(1)(d), 67(1)(d)(iv) [*Discrimination Act 1991*].

78. *Anti-Discrimination Act 1977* (NSW), ss 49ZXA-49ZXC [*Anti-Discrimination Act 1977*].

79. See Facebook Community Standards, Encouraging Respectful Behaviour, Hate Speech, online: www.facebook.com/communitystandards#. See also Facebook Statement of Rights and Responsibilities, s 3.7, online: www.facebook.com/legal/terms.

80. *Anti-Discrimination Act 1998* (Tas), ss 16(ea), 17(1)(ea) [*Anti-Discrimination Act 1998*].

81. *Anti-Discrimination Act 1977*, *supra* note 78 at ss 38R, 38S, 38T.

asexual (e.g., Northern Ireland (UK),⁸² England and Wales (UK),⁸³ Tasmania (Australia),⁸⁴ Australian Capital Territory (Australia)⁸⁵), while others name only homosexuality (e.g., New South Wales (Australia)⁸⁶). These laws are not limited to incitement to hatred laws, but the question of whether to protect sub-groups rather than general characteristics is as germane for incitement to hatred laws as for any hate speech laws. One relevant feasibility issue is the heavy evidential burden placed on legislators in identifying various sub-groups and determining how they compare with each other in terms of the relevant social evils. Processes of commissioning, receiving and digesting relevant research evidence could delay legislation for years and in some cases may even kick it into the long grass, but in the meantime the stirring up of hatred goes on unabated. Furthermore, facts about how different sub-groups compare to one another in relation to the relevant social evils are not static but change over time as circumstances change. And so if authorities want the body of law to reflect the fluidity of facts on the ground, there may be an impetus to create new stirring up hatred offences, then repeal them, then re-enact them, then repeal them, and so on, over relatively short periods of time. Given that parliamentary time is an extremely scarce and valuable resource, there would be considerable opportunity costs associated with adopting a highly fluid legislative programme. The net result might be that governments refrain from enacting any incitement to hatred laws for fear of committing themselves to the need for unlimited future changes to the law. In addition, it is surely a desideratum of any desirable legislative programme that the public will be able to keep up with changes so that at any given time they can predict or foresee, with a reasonable degree of certainty, whether or not their conduct is likely to be deemed a criminal offence by the police and the courts. In that sense a hokey cokey legislative programme may not be conducive to the rule of law and the values of publicity and predictability. Moreover, the splintering of the scope of incitement to hatred laws in terms of sub-groups may also pose not insignificant technical challenges for authorities in terms of drafting and enforcing these laws, such as finding workable legal definitions of the relevant sub-groups that can be rationally applied to concrete cases.

The aforementioned problems suggest, I think, at least two important things about how the functional approach relates to other approaches. The first is that other approaches will be indispensable in circumstances where different functional arguments point in different directions concerning whether or not to include a given characteristic under the scope of incitement to hatred laws. The second is that other approaches are likely to be highly relevant in deciding just how far to pursue the logic of given functional arguments concerning the proper scope of incitement to hatred laws, such as whether to frame these laws in terms

82. *The Public Order (Northern Ireland) Order 1987*, SI 1987/463 (NI 7), ss 8-13.

83. *Public Order Act 1986*, *supra* note 39, ss 29AB, 29JA, as amended by the *Criminal Justice and Immigration Act 2008*, *supra* note 63, and the *Marriage (Same Sex Couples) Act 2013*, *supra* note 63.

84. *Anti-Discrimination Act 1998*, *supra* note 80 at ss 16, 17(1)(c), 19(c).

85. *Discrimination Act 1991*, *supra* note 77 at ss 66(1)(b), 67(1)(d)(ii).

86. *Anti-Discrimination Act 1977*, *supra* note 78 at ss 49ZS, 49ZT, 49ZTA.

of broad categories like race, religion, sexual orientation, capacity, gender identity, and so on, or in terms of particular sub-categories like black, Muslim, gay, disabled, transgender, and so on. What these two problems do not suggest, however, is that the functional approach should play no role whatsoever in determining the proper scope of incitement to hatred laws.

II. Democratic specification

A final approach to specifying the proper scope of incitement to hatred laws places an emphasis on whatever scope best meets the special demands of democracy. This could take different forms, of course, depending on how these special demands are understood. One possibility is that it could be a matter of looking at whether or not certain groups in a democracy, perhaps because they are historically oppressed, have made demands for protection consistently over a prolonged period of time and for good reason or with justifiable cause. And so, when leaders within the UK Muslim community put requests to successive Home Secretaries for an extension of the existing stirring up hatred offences to cover Muslims, and did so consistently over a prolonged period of time and with justifiable cause, in the end it was fitting, democratically speaking, to introduce a new stirring up religious hatred offence (so the argument goes). Consider the words of the then Home Secretary, David Blunkett MP, who in 2001 sought (but failed at that time) to introduce a new stirring up religious hatred offence through the *Anti-Terrorism, Crime and Security Bill*.⁸⁷

The debate on this subject has been going on for a long time. We were approached by leaders of the Muslim community—it was a representative leadership group—who thought that it was only right, fair and protective to include religion with race in terms of avoiding incitement to hate using the Public Order Act 1986. I considered that and decided that their point was fair and reasonable.⁸⁸

There again, if democracy is about instituting rules that serve the wishes of the people who must live under the dominion of those rules, then it is unclear why any particular constituencies of society should have a privileged position. The democratically expressed wishes of hate speakers are as much a part of democracy as the wishes of those who are the objects of their hate speech. People who think it is their right to stir up hatred against persons identified by transgender identity, for example, are likely to demand that transgender identity is not treated as a protected characteristic for the purposes of incitement to hatred laws, whereas at least some of the people with a transgender identity against whom hatred is being stirred up are likely to demand the opposite. Faced with these conflicting wishes, authorities may find it difficult to justify ranking the wishes of hate speakers above the wishes of people who are the subject of hate speech but also *vice versa*, even if they accept the principle that a democratic society is one whose rules serve the wishes of the people.

87. Bill 49, 2001-2002 sess, 2001.

88. UK, HC, *House of Commons Debates*, vol 375, col 34 (19 November 2001).

Perhaps instead we should focus on the idea of democracy as a society organised according to rules that it gives to itself via systems of majority voting. In the case of representative democracy, this would involve a society electing political representatives to vote on the proper scope of incitement to hatred laws.⁸⁹ So, for example, in 2004 the two main opposition parties in the UK made it clear to the Labour government that they would not support the introduction of a new stirring up religious hatred offence via the *Serious Organised Crime and Police Bill*.⁹⁰ The government had to wait until the 2005 general election to cement its democratic mandate before introducing the bespoke *Racial and Religious Hatred Bill*.⁹¹ Viewing the “Who?” question in these terms might enable a reformulation of some of the distinctions introduced in Part 1. Thus, the distinction between unchosen and chosen characteristics could be reformulated not as a metaphysical or sociological distinction but rather as a democratic distinction. In short, the *real* distinction is between characteristics that a society or its democratically elected representatives does or does not decide, on a majoritarian basis, to treat as unchosen (quite apart from whether or not they are in fact unchosen) for the purposes of framing the scope of incitement to hatred laws.⁹²

Yet there are also some defects in this way of approaching the proper scope of incitement to hatred laws that may render it implausible as the sole approach. No society, and no body of elected representatives, will be unanimous about which, if any, characteristics it wants to protect under incitement to hatred laws and which it does not. At best a majority of democratic representatives will be reflecting a majority held view in society, and sometimes not even that. Moreover, what if the majority acts in extreme ways that wholly ignore the rights of the minority?

Part of the value of a constitutional democracy (it is often assumed) is that it provides checks and balances against potentially extreme results of majoritarian democratic decision-making. For the time being, the *Human Rights Act 1998*,⁹³ and with it the European Convention of Human Rights (ECHR) and the jurisprudence of the European Court of Human Rights (ECtHR), provides checks and balances against the decisions of the UK parliament *vis-à-vis* citizens' basic human rights, including in the area of criminal law. If, post-Brexit, the UK government opts to repeal the *Human Rights Act 1998* in favour of a British bill of rights, then this bill of rights will provide the sorts of rights-based constraints upon which the legitimacy of democratic government depends (it might be argued). Either way, arguably rights-based democracy-protecting and democracy-constituting checks and balances as well as majoritarian democratic decision-making should be key factors in determining the proper scope of incitement to hatred laws under the present democratic approach. Of course, this could also mean that the present approach is a subset of the functional specification approach. This is not to say it is necessarily

^{89.} Cf Erik Bleich, *The Freedom to Be Racist? How the United States and Europe Struggle to Preserve Freedom and Combat Racism* (Oxford: Oxford University Press, 2011) at 12-13.

^{90.} Bill 5, 2004-2005 sess, 2004.

^{91.} *Supra* note 55.

^{92.} Cf Robert C Post, “Racist Speech, Democracy, and the First Amendment” (1991) 32:2 Wm & Mary L Rev 267 at 295 [Post, “Racist Speech”].

^{93.} (UK), c 42 [*Human Rights Act*].

wrong, of course. At any rate, what implications will this have for questions of scope? For one thing, it could mean that there is an impetus for the scope of hate speech laws to harmonise with the scope of human rights. The *Human Rights Act 1998* gives effect to the ECHR, Art 14 of which makes it clear that ‘[t]he enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status’. And Art 10(2) states that the exercise of the right to freedom of expression protected under Art 10(1) may be restricted by laws that ‘are necessary in a democratic society’.⁹⁴ Together these might suggest incitement to hatred laws whose scope includes political opinion as a protected characteristic, for example. A British bill of rights might also incorporate the idea that the right to freedom of expression is fundamental but not absolute, and may be limited for the sake of democratic values. In terms of the relevant protected characteristics, it is not inconceivable that a new British bill of rights could borrow the list of characteristics already found in UK anti-discrimination law, namely, age, disability, gender reassignment, race, religion or belief, sex, sexual orientation.

The key point here is that the issue of what, if any, limitations on freedom of expression are necessary in a democratic society is contested, and this will inevitably have implications for the “Who?” question. Both Robert Post and James Weinstein have sought to recast the relationship between the right to freedom of expression and the ideal of democratic self-government by highlighting an earlier stage in the democratic process. On their approach, what really matters is the rights of individuals to participate in the formation of public opinion upon which the familiar mechanisms of democratic decision-making are based. They argue that any hate speech laws which restrict public discourse cannot be warranted regardless of their scope.⁹⁵ However, arguably some hate speech laws, even those that restrict the opportunities of hate speakers to choose their preferred modes of speech, delivery and timing when participating in public discourse, can be *pro tanto* warranted if they operate for the sake of ensuring that all citizens enjoy real opportunities to participate in public discourse. The basic concern here is that, if left unchecked, certain forms of hate speech can deter or inhibit members of targeted groups from functioning as ordinary deliberative democrats. What is at stake is a sort of deliberative exclusion in which, out of fear for their personal safety or livelihood or as a result of an impaired sense of their status, some,

94. ECHR, online: www.echr.coe.int/Documents/Convention_ENG.pdf.

95. See, e.g., Post, “Racist Speech”, *supra* note 92; Robert C Post, “Managing Deliberation: The Quandary of Democratic Dialogue” (1993) 103:4 Ethics 654; Robert C Post, “Hate Speech” in Ivan Hare & James Weinstein, eds, *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009) 123; Robert C Post, “Participatory Democracy and Free Speech” (2011) 97:3 Va L Rev 477; “Interview with Robert Post” in Michael Herz & Peter Molnar, eds, *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge: Cambridge University Press, 2012) 11. See also, e.g., James Weinstein, “Hate Speech, Viewpoint Neutrality, and the American Concept of Democracy” in Thomas R Hensley, ed, *The Boundaries of Freedom of Expression and Order in a Democratic Society* (Kent: Kent State University Press, 2001) 146; James Weinstein, “Extreme Speech, Public Order, and Democracy: Lessons from the Masses” in Ivan Hare & James Weinstein, eds, *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009) 23; James Weinstein, “Participatory Democracy as the Central Value of American Free Speech Doctrine” (2011) 97:3 Va L Rev 491.

perhaps many, victims of hate speech might (i) refrain from participating in the formation of public opinion; (ii) adapt their expressed preferences in order to fit their reduced circumstances; or (iii) discover that even when they do decide to speak up what they say falls on deaf ears because of the low opinion that others have of them.⁹⁶ Part 3A of the *Public Order Act 1986*⁹⁷ sets out offences of using ‘threatening words or behaviour’ to stir up hatred against people on grounds of their religion or sexual orientation. At first glance, this restrictive wording is hard to understand. If the concern is ultimately with the likelihood of hatred being stirred up, why restrict the offenses to only threatening words or behavior? Why not also include insulting, derogatory, stigmatising, or defamatory words? One possible answer is that these offences are designed to protect the subjective element of insecurity, people’s sense of fear and insecurity. And so, these stirring up offences are designed to deter the sort of incitement to hatred that not merely creates a climate of hatred but also a climate of fear and insecurity. The current line of argument looks to the knock on effects of a climate of fear and insecurity measured in terms of a loss of real opportunities to participate in public discourse, as in, opportunities that are free from fear. Once again, it is not hard to see how these sorts of arguments might be plausibly applied not merely to what are the currently protected characteristics of race, religion and sexual orientation under the stirring up hatred offences in England and Wales but also to disability and gender identity in the event that the stirring up of hatred also creates a climate of fear and insecurity that silences members of these other historically victimised and still vulnerable groups or classes of people.

Notwithstanding these points, it might be argued that I have missed the true purpose of protecting fundamental rights, such as the right to freedom of expression, within a constitutional democracy. It is, as Ronald Dworkin argues, to ensure the democratic legitimacy of all other laws.⁹⁸ To explain, Dworkin argues that if we, as a democratic society, introduce “upstream” hate speech regulations and thereby “intervene too soon in the process through which collective opinion is formed”, then “we spoil the only democratic justification we have for insisting that everyone obey [downstream] laws”,⁹⁹ including antidiscrimination laws that also protect the very groups claiming protection under hate speech regulations. Consider an example due to Waldron.¹⁰⁰ Suppose an English landlord discriminates against families of South Asian descent in a way that is prohibited by English antidiscrimination laws. At the same time, English laws banning the stirring up of racial hatred prevent the landlord from using threatening, abusive or insulting words or behaviour with either the intention or likelihood of stirring

96. For a list of references to supporting literature, see Brown, “*Hate Speech*”, *supra* note 2 at 198.

97. *Supra* note 39, as amended by the *Racial and Religious Hatred Act 2006*, *supra* note 58, and the *Criminal Justice and Immigration Act 2008*, *supra* note 63.

98. See, e.g., Ronald Dworkin, “Foreword” in Ivan Hare & James Weinstein, eds, *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009) at v [Dworkin, “Foreword”]; Ronald Dworkin, “Reply to Jeremy Waldron” in Michael Herz & Peter Molnar, eds, *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge: Cambridge University Press, 2012) 341.

99. Dworkin, “Foreword”, *supra* note 98 at viii.

100. Waldron, *supra* note 33 at 1643.

up hatred against Pakistanis defined as a racial, ethnic or national group. This means the Landlord is restricted from participating, in the ways he might prefer to participate, in public discourse and the formation of public opinion about the antidiscrimination laws he objects to. On Dworkin's analysis (pace Waldron¹⁰¹), this has a significant negative impact on the legitimacy of those same antidiscrimination laws. How so? One possible explanation invokes Justice Scalia's metaphor of expecting some people to follow the Marquis of Queensberry rules whilst their opponents fight freestyle.¹⁰² The stirring up racial hatred offences prevent the landlord from using threatening, abusive or insulting words or behaviour with either the intention or likelihood of stirring up hatred against Pakistanis or against supporters of the antidiscrimination laws identified racially. But at the same time these offences do not currently prevent Pakistanis or any defenders of the antidiscrimination laws from using threatening, abusive or insulting words or behaviour with either the intention or likelihood of stirring up hatred against landlords identified as a group or class of persons *in virtue of being landlords*.¹⁰³

However, notice that this particular explanation of why incitement to hatred laws in England and Wales (and other countries) damage the legitimacy of so-called downstream laws makes an assumption that no sensible incitement to hatred laws would include among the relevant protected characteristics the characteristic of landlord along with the more conventional characteristics race, ethnicity, nationality, religion, sexual orientation, disability, gender identity, and so on. But this assumption may not be as sound as it first appears, especially if there are solid philosophical reasons for the inclusion of the broader characteristic *profession* as a protected characteristic. Thus, as I discussed in Part 1, suppose one believes that the proper scope of incitement to hatred laws should be determined partly on the basis of whichever characteristics, social identities or statuses tend to be an integral feature of the subjective personal identities of the people who possess them. It might be argued that for many people profession is something they accept or adopt as a central part of who they really are.¹⁰⁴ At any rate, one cannot simply discount out of hand the idea that hate speech laws should also include profession as a protected characteristic without appealing to a comprehensive theory of the proper scope of hate speech laws, that is, without having a good answer to the “Who?” question that takes into consideration various different approaches to answering that question.

III. Conclusion

Across Parts 1 and 2 I have outlined and evaluated five main approaches to specifying the proper scope of incitement to hatred laws. In doing so I have suggested

101. *Ibid* at 1646.

102. *R A V v City of St. Paul* (1992) 505 US 377 at 392 (involving the application of the City of St. Paul's Bias-Motivated Crime Ordinance to a case of cross burning on the property of an African American family).

103. Cf James Weinstein, “Hate Speech Bans, Democracy and Political Legitimacy” *Constitutional Commentary* [forthcoming in 2017].

104. See Brown, “Who”, *supra* note 1 at 309–11.

strengths as well as weaknesses in each of the approaches. Perhaps not surprisingly, then, my overall conclusion is that the correct answer to the question ‘How do we decide who should be protected by incitement to hatred laws?’ is this: ‘With a combination of all five approaches.’ No specification would be adequate, I believe, that appealed exclusively to any one of these approaches. Conversely, a specification would be lacking or incomplete if it disregarded entirely any one of these approaches. To be specific, I think one would be hard pressed to say that parity and consistency alone can determine the proper scope of incitement to hatred laws, but, conversely, it would be too big a jump to conclude that consistency or parity have no relevance to these questions whatsoever. When legislators in England and Wales have sought to create new stirring up hatred offences they have not, at least not since 1965, done so with a *tabula rasa*. The very fact that some stirring up hatred offences already exist necessitates, to some extent at least, arguments based on principles of parity and consistency—principles that may have inherent as well as instrumental value. By the same token, the idea that authorities could create a new offence purely for the sake of parity, say, even though there are no actual occurrences of prosecutable cases, may sound a little absurd to some people (“political correctness gone mad”). Then again, practical specification only asks if there is a phenomenon of stirring up hatred on grounds of a given characteristic and if, practically speaking, there is something that could be done about it through the creation of new offences. What it cannot tell us is whether or not it would be right to introduce such new offences. Likewise, formal specifications could never tell the whole story about the proper scope of incitement to hatred laws partly because understanding the full significance of formal qualities often depends on thinking about the underlying function or purpose of these laws. On the other hand, arguably formal distinctions add sophistication to the way we compare and contrast different characteristics, especially when applying the principle of parity (treating like groups alike), because they help us to know the deeper and myriad ways characteristics can be alike or unalike. As for functional specifications, these use the underlying or real function or purpose of incitement to hatred laws in order to shape the scope of such laws. Nevertheless, from the mere fact that creating a new stirring up hatred offence for a given characteristic would serve the basic function or purpose of incitement to hatred laws in general, one cannot safely assume that this would be the proper thing to do, bearing in mind practical and formal considerations. That being said, surely it would be irrational, impossible even, to completely ignore or set aside functional considerations, to detach the “Who?” question from the “Why?” question. Finally, democratic specifications speak to how the scope of incitement to hatred laws may or may not meet the special demands of democracy. But it is not hard to think of reasons why allowing majoritarian decision-making alone to determine scope may be unwise. For its part, the idea of constitutional democracy may well rely on ideas of what is necessary in a democratic society that in turn rely on functional specifications. On the other hand, any approach that rejected the special demands of democracy altogether would seem to ignore, violate even, what lends special value and even legitimacy to the very systems of government in which incitement to hatred laws operate.

But if the best approach is one that combines all five approaches, how is that combination to be done? What does it look like? My view is that the most likely combination is not going to be a strict hierarchy or lexical priority. I think it extremely unlikely, in other words, that consistency, practical, formal, functional and democratic specifications can be put into an obvious or self-evident single rank order, so that whichever approach is ranked first determines the proper scope, the second breaks any ties remaining from applying the first, the third breaks any ties remaining from applying the second, and so on. For example, it might be tempting to think that functional specification is lexically prior given the role it can often play in motivating, limiting and in some instances substantiating arguments within the other four approaches. Yet it would be quite dogmatic, I think, to suppose that the underlying or real function or purpose of incitement to hatred laws could in theory settle the issue of the scope of such laws in a way that effectively trumps anything of a contradictory nature that the consistency, practical, formal and democratic approaches may have to say.

I also think that any talk of achieving an optimum “balance” or “trade-off” between the five approaches assumes something about the commensurability of these approaches that is actually belied by their plural and heterogeneous nature. So if not balancing then what? I have argued elsewhere that the question ‘Can hate speech laws be warranted?’ must be settled through compromise,¹⁰⁵ and I am inclined to say that the same can, and should, be said in response to the question ‘What is the right combination of approaches to determining the proper scope of hate speech laws?’. In other words, the best strategy is one that combines together all five approaches in reasonable ways given the fact that sometimes it will be necessary to compromise some approaches for others if it is not possible to honour all approaches at the same time: that is, if each approach points to a slightly different scope.

Clearly context will also be important for how any reasonable compromise of approaches plays out. The main point I wish to make about context is that many of the considerations which I have put forward in Parts 1 and 2 will take on different dimensions depending on the social, legal and political conditions of the given country. In England and Wales, for example, some people argue that in order to justify the introduction of new stirring up hatred offences there has to be a very serious threat to public order or security. This is a functional approach to scope specification. Yet the threat posed by incitement to hatred on grounds of disability, gender or transgender identity (so they argue) is not on a par with the threat associated with race or religion. Our society has experienced religious wars and race riots, after all, but nothing comparable in terms of these other characteristics. Then again, the justification required for new stirring up hatred offences need not be exactly the same as the justification required for existing stirring up hatred offences. Indeed, in the case of sexual orientation nobody seriously suggested that the UK was on the brink of a kind of sexual orientation war, *hetero versus homo*. What matters is not public disorder in the literal sense of

¹⁰⁵ Cf Brown, “*Hate Speech*”, *supra* note 2 at ch 10.

race riots. That certain groups can be left feeling insecure as a result of incitement to hatred may be enough. And the feelings of fear and insecurity produced by different forms of incitement to hatred might be more similar than we think, not least in their silencing effects. Nevertheless, in other contexts the underlying or real function or purpose of incitement to hatred laws might be very different, and so the arguments around scope could play out very differently. In China, for instance, the underlying or real function or purpose of hate speech laws is to tackle speech which can undermine national harmony in the sense of causing friction between national minorities. Here the context is a substantially nondemocratic governmental regime covering a geographically huge state-nation made up of provinces some of which contain significant national minority groups that for some time have been agitating for greater political autonomy and even full independence. Given this underlying or real function or purpose, it may not be all that surprising that relevant hate speech laws cover only nationality or nationality minority status.¹⁰⁶ In this context the constitutional democracy approach might take on an even greater urgency and importance, in terms of widening the scope of hate speech laws.

Notwithstanding these observations about the importance of context, I will now attempt to draw some tentative conclusions of a general kind about which particular characteristics, identities or statuses ought to be protected under incitement to hatred laws informed by all five approaches. First, I believe that there is something approximating a core set of characteristics that tend to be supported to a *significant* extent by *most*, if not all, of the five main approaches (under their best interpretations). Not only do they crop up more often in bodies of law, legal literature and parliamentary debates, but it does seem that a strong case can be made, based on the five approaches, for lifting them into the position, if they are not there already, of protected characteristics. These are: race, ethnicity, nationality, regional nationality, citizenship status, religion, sexual orientation, gender (or gender identity), and disability. That being said, even for this core set of characteristics, there remain dilemmas about how their basic elements are to be defined and what their boundaries are. And, once again, resolving these dilemmas relies on some reasonable combination of the five approaches. Second, I also believe that there is a much longer list of marginal or borderline characteristics that crop up more sporadically in the aforementioned arenas and arguably are supported only to *some* extent by only *some* of the five approaches. These are: age, age performance, education, employment status, gender performance,¹⁰⁷ language status, marital status, medical status, parental status, personality traits or types, physical appearance, political beliefs, activities, or affiliations, pregnancy, profession, sexual preference, social status, and war record. Some of the characteristics are relegated to this list of borderline protectable characteristics because of practical considerations, such as lack of substantial evidence to demonstrate that there is

106. *Criminal Law of the People's Republic of China*, 1979/83, arts 249, 250 (amended in 1997 and 2009).

107. It might be tempting to subsume gender performance under the core characteristic of gender, but doing so belies differences between the two that might come to the fore on some of the aforementioned approaches.

currently a problem of stirring up hatred related to them, whereas some are on the list in virtue of relevant formal and functional considerations. Most or all of them appear on the list in virtue of concerns over parity or even consistency.

Note, I call these “tentative” conclusions because for reasons of space I have only been able to present and sketch out with some illustrations a general framework or methodology for answering the “Who?” question. Although I have also put forward some substantive arguments about particular characteristics when illustrating and critically examining each of the five approaches to specification, I have not attempted to provide anything close to a comprehensive and definitive set of arguments for each one of the numerous characteristics. This would obviously require a much longer treatment. My two-part article has been designed merely to give a wide-angle overview of the landscape. Therefore, my conclusions should be read as working hypotheses based on finite research.

Now I also made it clear in Part 1 that my observations about the proper scope of hate speech laws are intended to relate specifically to incitement to hatred laws and that it is possible that some different conclusions would follow if the focus is shifted to other types of hate speech law. I stand by that claim. But, at the same time, I do not suppose that what I have said here is entirely idiosyncratic to this particular cluster of hate speech laws. It seems likely that some of the substantive arguments as well as the five approaches will be echoed in a discussion one could have about the proper scope of other types of hate speech law. In order to motivate this conjecture about partial overlap (but not identity) of plausible answers to the “Who?” question across different types of hate speech law, I shall take this opportunity to make some condensed remarks about one other variety of hate speech law. Consider laws that disallow, either directly or indirectly, the use of targeted insults, slurs, derogatory epithets or other abusive comments against individuals who are picked out by protected characteristics, especially in face-to-face encounters but also online.¹⁰⁸ Here again I believe that issues of consistency and parity are likely to be important. Consider the fact that ss 31 and 32 of the *Crime and Disorder Act 1998*¹⁰⁹ set out discrete racially or religiously aggravated public order and harassment offences. One such public order offence is racially or religiously aggravated intentional harassment, alarm or distress: namely, the offence of, with intent to cause harassment, alarm or distress and aggravated by racial or religious hostility, using threatening, abusive or insulting words or behaviour, or displaying writing, signs or other visible representation that are threatening, abusive or insulting, and thereby causing harassment, alarm or distress.¹¹⁰ One recent high-profile case was *R v. Stacey*.¹¹¹ By contrast, there are no discrete aggravated public order or harassment offences covering the same conduct when it is

108. See, e.g., Brown, “Hate Speech”, *supra* note 2 at 23–26.

109. (UK), c 37 [*Crime and Disorder Act 1998*], as amended by the *Anti-Terrorism, Crime and Security Act 2001* (UK), c 24 [*Anti-Terrorism, Crime and Security Act 2001*].

110. *Crime and Disorder Act 1998*, *supra* note 109, s 31(1)(b).

111. (2012), No. A20120033, Swansea Crown Ct, 30 March (involving messages sent by the defendant on the Twitter Internet messaging service in response to, and as part of a heated discussion about, a life-threatening heart attack suffered by a black footballer during a live broadcast match).

aggravated by hostility toward victims based on their sexual orientation, disability, or transgender identity. Instead, such cases would be covered by generic hate crime or sentence enhancement provisions, that is, ss 145 and 146 of the *Criminal Justice Act 2003*,¹¹² which give magistrates and judges the power to determine the seriousness of the offending, with hostility toward people based on race, religion, sexual orientation, disability, or transgender identity as an aggravated factor. Why does this difference matter? It matters because the sentencing structure is different depending on the offence, and this in turn means different treatment for similar conduct and different levels of protection for similar groups. So, for example, the discrete racially or religiously aggravated offence of intentional harassment, alarm or distress set out in s 31(1)(b) of the *Crime and Disorder Act 1998*¹¹³ can be prosecuted either as a summary offence or as an indictable offence in crown court. In the latter instance, a person found guilty of an offence on indictment can be liable to imprisonment for up to two years.¹¹⁴ However, the equivalent basic offence of intentional harassment, alarm or distress set out in s 4A of the *Public Order Act 1986*¹¹⁵ is only a summary offence handled in magistrates courts. A person found guilty of an offence on summary conviction can only be liable to imprisonment for a term not exceeding six months.¹¹⁶ Therefore, even though a magistrate has the power under sentence enhancement provisions, that is, ss 145 and 146 of the *Criminal Justice Act 2003*, to take into account hostility toward victims based on disability, sexual orientation or transgender identity as an aggravating factor in judging the seriousness of a basic offence of intentional harassment, alarm or distress, a magistrate must operate within the aforementioned sentencing framework, meaning that he or she cannot exceed the maximum sentence of six months imprisonment for the basic offence.¹¹⁷

A full inquiry as to the proper scope of insult-focused hate speech laws must also inevitably turn to questions of underlying or real function or purpose. One obvious place to start is with harm. In the case of hate speech that takes the form of directly targeting insults, slurs, derogatory epithets or other abusive comments at individuals who are picked out by protected characteristics, especially in face-to-face encounters but also online, the harm may compromise many of the social evils discussed above but it can also take the form of immediate or short-term emotional distress but also medium- to long-term psychological and physiological ill-health.¹¹⁸ Once again, the logic of such functional arguments may point in the direction of shaping the scope of hate speech laws around particular sub-groups. Mari Matsuda has argued in relation to the US, for example, that laws

^{112.} (UK), c 44 [*Criminal Justice Act 2003*], as amended by the *Legal Aid, Sentencing and Punishment of Offenders Act 2012* (UK), c 10, s 65 [*Legal Aid, Sentencing and Punishment of Offenders Act 2012*].

^{113.} *Supra* note 109.

^{114.} *Supra* note 109, ss 31(1)(b), 31(4), as amended by the *Anti-Terrorism, Crime and Security Act 2001*, *supra* note 109.

^{115.} *Supra* note 39.

^{116.} *Supra* note 39, s 4A(4).

^{117.} *Supra* note 112, as amended by the *Legal Aid, Sentencing and Punishment of Offenders Act 2012*, *supra* note 112.

^{118.} See, e.g., Brown, “*Hate Speech*”, *supra* note 2 at 51–56.

criminalising or allowing civil redress for insults should be limited to protecting historically oppressed, victimised, persecuted or systematically discriminated against or disadvantaged groups, such as racial minorities.¹¹⁹ This speaks to the idea that the emotional and psychological toll of racist insults are of a special or *sui generis* quality, magnitude, and likelihood because of the history of slavery and other forms of racial oppression.¹²⁰ Kent Greenawalt, however, argues that, depending on the country context, group-identity-based insults or epithets may be especially hurtful or distressing not just for groups defined by race but also for groups defined by religion, gender, and sexual orientation.¹²¹ Mirroring existing scholarly arguments concerning the special nature and severity of emotional distress caused by racist insults due in large part to the history of black oppression, it is argued that insults directed at people based on their sexual orientation have a special power to cause psychological hurt or emotional distress due in large part to the history of homophobic discrimination and violence.¹²² Perhaps similar arguments could be made about the particular emotional distress and psychological damage caused by identity-based insults when they are directly targeted at individuals on grounds of their possessing disabilities or particular gender identities once again reflecting the history of discrimination and violence against members of such groups. Then again, it may seem wrong to make determinations of scope solely on the basis of purely functional arguments. Alon Harel, for example, argues that basing decision about the proper scope of insult-focused hate speech laws exclusively on the likelihood of hate speech against a given group of people causing social tensions or public disturbances would be “inherently demeaning”.¹²³ According to Harel, when it comes to laws disallowing identity-based insults or epithets such as “nigger” or “kike”, to treat anti-black insults differently to anti-Semitic insults “simply because the racial hatred directed at Jews is less likely to lead to racial friction”¹²⁴ would be to violate the abstract principle of treating citizens as equals.

In addition to this, there are bound to be balancing-based considerations in deciding the scope of insult-focused hate speech laws. One consideration might be

¹¹⁹. See, e.g., Matsuda, *supra* note 7 at 2357.

¹²⁰. See, e.g., Richard Delgado, “Words That Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling” (1982) 17:1 Harv CR-CLL Rev 133 at 179-81 [Delgado]; Kretzmer, *supra* note 4 at 458; Charles R Lawrence III, “If He Hollers Let Him Go: Regulating Racist Speech on Campus” (1990) 3 Duke LJ 431 at 458-59; Akhil Reed Amar, “The Case of the Missing Amendments: R.A.V. v. City of St. Paul” (1992) 106:1 Harv L Rev 124 at 125-26; Joshua Cohen, “Freedom of Expression” (1993) 22:3 Phil & Pub Aff 207 at 255-56; Stephen J Heyman, “Spheres of Autonomy: Reforming the Content Neutrality Doctrine in First Amendment Jurisprudence” (2002) 10:3 Wm & Mary Bill Rts J 647 at 690-91, 713-14; Stephen J Heyman, *Free Speech and Human Dignity* (New Haven: Yale University Press, 2008) at 273, n 26.

¹²¹. Greenawalt, “Speech”, *supra* note 44 at 147-48, 156, n 26; “Fighting Words”, *supra* note 44 at 55.

¹²². See, e.g., Toni M Massaro, “Equality and Freedom of Expression: The Hate Speech Dilemma” (1991) 32:2 Wm & Mary L Rev 211 at 223; Major, *supra* note 11 at 227; Cohen, “More Censorship”, *supra* note 8 at 74; Leets, *supra* note 51 at 350, 356.

¹²³. Alon Harel, “Bigotry, Pornography, and the First Amendment: A Theory of Unprotected Speech” (1992) 65:4 S Cal L Rev 1887 at 1906 [Harel].

¹²⁴. *Ibid.*

that insults pertaining to disability and gender identity, for example, are so commonplace and embedded in our ordinary language that to disallow such speech would be to disallow a high volume of speech, thus requiring unprecedented levels of censorship and curtailment of free speech.¹²⁵ However, to this argument it can be countered that at one time insults pertaining to race and religion were also a ubiquitous feature of ordinary language but that was not a legitimate basis for failing to combat them. Moreover, the mere fact that a law tends to render illegal a high volume of everyday or commonplace conduct is not always considered a decisive reason for repealing that law or for not enacting it in the first place, such as in the case of driving in excess of speed limits.¹²⁶

Furthermore, some of what can be said about the proper scope of insult-focused hate speech laws might appear to reflect formal specification but is in fact functional specification. Richard Delgado, for example puts a slightly different spin on the moral significance of immutability, transforming it into a functional argument. According to Delgado, “[t]he psychological harms caused by racial stigmatization are often much more severe than those created by other stereotyping actions” precisely because “membership in a racial minority can be considered neither self-induced, like alcoholism or prostitution, nor alterable”.¹²⁷ Racist hate speech (according to Delgado) can cause double misery: misery in being subjected to hate speech in virtue of one’s race; and misery in the knowledge that one is trapped, that there is nothing one can do to change the attribute in virtue of which one is being so subjected. If this is true, then surely similar arguments could be made about many other characteristics, social identities or statuses.

Finally, I wish to end the article with two general remarks about the wider implications of my examination of the “Who?” question. I have used a combination of consistency, practical, formal, functional and democratic specification to generate both a relatively short list of core protected characteristics and a much longer list of borderline protectable characteristics. However, I must now address a potential objection to incitement to hatred laws *tout court*, namely, that to create *any* list of protected characteristics, and especially a long list of borderline characteristics, creates a slippery slope to a body of legislation that would be, in no uncertain terms, either an unmitigated disaster for, or a gross violation of, the right to freedom of expression. I believe the foregoing research provides one possible line of response to the slippery slope objection. The objection relies on the fact that once one characteristic is granted protection through hate speech laws, there are no relevant considerations on which to deny similar protection to other characteristics and associated groups of people. The research demonstrates that there are in fact at least five main families of considerations which could each potentially provide stopping points.

Second, I believe that the foregoing research may well have a wider significance beyond research on hate speech law. Political theorists who work on and

^{125.} See, e.g., Heinze, “Cumulative Jurisprudence”, *supra* note 47 at 274-75, 278-79.

^{126.} Cf Weston-Scheuber, *supra* note 32 at 149.

^{127.} Delgado, *supra* note 120 at 136.

with liberal theories of equality of opportunity as well as political theorists who work on and with theories of recognition, the politics of difference, multiculturalism, and intercultural dialogue, may benefit from the research by gaining a better understanding of what it is about groups or classes of persons that makes them eligible for not merely protections against certain forms of mistreatment (e.g., hate speech) but also exemptions from general laws, special accommodations when it comes to access to public goods and the provision of public services, and respect in matters of intercultural dialogue. The point is that although I have focused on the scope of hate speech laws, it may be that the five approaches I have outlined may be applicable to questions about the proper scope of protections in anti-discrimination and employment law, exemptions in road safety law, accommodation practices in built environment and urban planning policy, and rights to respectful communication in intercultural dialogue regulation, to name but a few examples.