Ce document est librement téléchargeable sur la page web du CDPC : http://www.u-paris2.fr/CDPC0/0/fiche pagelibre/.
Pour le citer : PHILLIPSON Gavin, « Hate Speech Laws : What They Should and Shouln't Try to do", Conférence-débat du CDPC, 9 avril 2015, URL, date de consultation.

HATE SPEECH LAWS: WHAT THEY SHOULD AND SHOULDN'T TRY TO DO

Professor Gavin Phillipson*

I. Introduction

Hate speech laws are enduringly controversial. In Europe & in Commonwealth counties, including Canada, Australia, South Africa and the UK, the position is taken that bans on hate speech are not only permissible under human rights standards, but actively *required* by them. The European Court of Human Rights has upheld numerous hate speech bans and even held that direct expression of racial hatred is not 'expression' for the purposes of Article 10(1) at all;[1] Article 20 of the International Covenant on Civil and Political Rights *requires* hate speech bans.[2] So in democracies other than the US it is now almost taken for granted that incitement to racial hatred should be proscribed; the question is how far to extend the protection – whether to include grounds of religious belief, sexual orientation, gender status, disability and so on. Meanwhile, many Islamic countries have lobbied for the extension or creation of worldwide blasphemy laws to cover Islam, so that the publication of the 'Danish cartoons', for example, would have been a crime. In sharp contrast, the US continues resolutely to uphold its stance that all hate speech bans – as well as blasphemy laws - violate the First Amendment.[3] Hence in the US, alone in the democratic world, intentional incitement to racial hatred is constitutionally protected speech. However, the divide is not simply one of the US v the Rest. There are scholars in the US sympathetic in principle at least to hate speech bans, while a number of European and Commonwealth scholars are influenced to some extent by the US arguments, and sharply critical of at least the scope, if not the existence of many hate speech bans in the world's democracies.

It's really this sharp divide that led to my interest in this issue and the approach I'm taking in my forthcoming book with Professor Heinze: *Debating Hate Speech*.[4] Living in a European democracy, you grow up taking for granted hate speech bans. Then you come across the US literature, which provides a series of very persuasive and passionate arguments as to why hate speech bans are both wrong as a matter of principle and ineffective – even counter-productive – in practice. This

* Professor of Law, Durham Law School, Durham University https://www.dur.ac.uk/law/staff/?id=5039
Unpublished draft; please do not cite without author's consent.
1 *Jersild v Denmark* (1994) 19 EHRR 1.
2 ICCPR Art 20: 'Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.' (The USA entered a reservation on this point).
3 Giving constitutional protection to freedom of speech; see e.g. *R.A.V. v. City of St. Paul,* 505 U.S. 377 (1992).
4 Hart Publishing, forthcoming 2016.

forces you to think again. I have concluded in the end that the US arguments about the *inefficacy* of existing hate speech laws are largely right – but are generally ignored in Europe and the Commonwealth, in which we tend too often simply to *assume* that hate speech bans must do some good. I have also concluded that the US arguments on principle are powerful ones and that many European hate speech laws rightly fall victim to them, through their lazy conflation of insult with intimidation, hurt religious sensibilities with hatred of peoples, offensiveness with dignitarian assault. I don't in the end agree with the First Amendment proposition that no speech can be banned on the basis of its viewpoint. But having said that I think the US arguments can be of great use to the rest of us. My view is that should be *harnessed*: not as a weapon to kill off hate speech bans completely, but as a surgeon's instrument, to cut away the flabby tissue surrounding pro-ban thinking, leaving a robust and healthy principle behind, a principled core. It is that core that I try to defend in this paper.

It is important to be clear what I am talking about when I use the term 'hate speech'. The key controversy does not apply to all of what we might think of as hate speech but only what American scholars tend to term 'contributions to public discourse'. Thus 'hate speech' for my purposes does *not* include any of the following: so-called 'fighting words' (that is, face-to-face threats or insults); targeted messages, e.g. emails, twitter messages, telephone calls, text messages; incitement to violence against particular individuals: 'let's go kill Phillipson', defamation of particular individuals or a small, identifiable group; private conversation[5]. The subject of this paper is therefore those laws that apply to speech addressed to the public at large – for example, uttered during public demonstrations, printed in newspapers or magazines, or posted on a blog, and which are about racial or other groups, not individuals.

II. Existing Hate Speech Laws: the many problems.

European hate speech bans have remained controversial not just because of the academic debate around them, but because of their breadth and the results. Many view the resultant jurisprudence as showing a conflation of the notion of hate speech with merely offensive speech, particularly in the field of religious belief – a conflation that opens the danger of supposed hate speech laws acting as proxy for blasphemy laws. Chief among the problems here has been the approach of the European Court of Human Rights which has, in a string of decisions that many regard as disastrous for free speech, including *Otto-Preminger v Austria*[6] and *IA v Turkey*,[7] upheld the rights

[5] All such actions other than the last will be subject to criminal or civil liability under other laws.
[6] (1995) 19 EHRR 34.
[7] Application no. 42571/98 (2005).

of governments to silence speech directed against the dominant religious group, simply on the basis that they find it gratuitously offensive to their religious feelings. And this has not even been about protecting vulnerable *minority* religious groups from vilification: the main cases have concerned Roman Catholicism in Austria, Anglicanism in the UK,[8] Islam in Turkey. The Strasbourg Court has even implausibly claimed that religiously offensive expression threatens people's ECHR right to religious freedom.[9] These findings, say many, have had nothing to do with upholding basic citizen dignity from attack; rather they have simply validated state control over the artistic and political treatment of major figures of dominant established religions. The same thing applies in many European countries. The case which the French government lost at Strasbourg, *Giniewski v France*[10] is a good example: this was a 'group defamation' case, in which a newspaper article made the argument that there was a causal connection between centuries of Catholic anti-Semitism and the Nazi Holocaust; the publishing director and the journalist were found criminally and civilly liable. This judgment, in my view, directly attacked the freedom to debate an important and controversial topic under the guise of upholding the 'honour' of Christians. Even Strasbourg found it to be an unwarranted restriction on freedom of expression.

These overly broad laws are then rightly vulnerable to powerful lines of US argument: that they allow the state to suppress speech it disapproves of – speech ridiculing or attacking the Catholic Church for example; that they impose a state orthodoxy on views that should be open to debate; that in punishing speakers with whom it disagrees it is engaging in thought control; that it is invading the moral autonomy of the speakers who are punished, thereby failing to treat them with the equal respect that the state must extend to all its citizens, as Ronald Dworkin has argued.[11]

There are further problems. Because so many hate speech laws are concerned with "insulting or abusive" speech, to use the terminology of the British racial hatred laws, US scholars rightly argue that the law ends up at least in part by enforcing matters of taste and decorum and is therefore inherently elitist in its impact:[12] the uneducated man, who resorts to crude language to express his unease about Muslim immigration, or his Christian disapproval of gay sex, is

[8] *Wingrove v United Kingdom* (1997) 24 EHRR 1.

[9] Protected under Art 9 ECHR. See *Otto Preminger* at [49]*:* "[the responsibilities of those exercising the Article 10 right include] an obligation to avoid as far as possible expressions that are gratuitously offensive to others and thus an infringement of their rights'; *IA v Turkey* (at [29])*,* in which the Court said that its basic task was 'weighing up the conflicting interests of the exercise of two fundamental freedoms, namely the right of the applicant to impart to the public his views on religious doctrine on the one hand and the right of others to respect for their freedom of thought, conscience and religion on the other hand'.

[10] (2007) 45 EHRR 23

[11] See e.g. 'Do we have a right to pornography?' in *A Matter of Principle* (1985).

[12] See Jim Weinstein's essay in I. Hare and J. Weinstein, *Extreme Speech and Democracy* (Oxford, OUP, 2009).

punished because he uses crude epithets, like 'Pakis' or 'poofs'. The intellectual, who writes books arguing for the inherently violent nature of Islam, or those who publish the Bible, which in Leviticus condemns gay sex as 'an abomination' (18:22), punishable by death (20:13) are untouched by law. There are intractable problems here. While passages like this in the Bible forcibly express the notion that the practice of homosexuality is sinful, the English High Court has found that a Christian speaker publicly calling on people to desist from gay relationships (arousing angry dissent in his listeners in the street) was using "insulting" words for the purposes of public order law[13] and could therefore be arrested and convicted of an offence. But what then about the Danish Cartoons – or similar cartoons of the Prophet Mohammed published by *Charlie Hebdo*? Aren't they experienced as insulting, as causing great distress? Many Muslims say yes. Many worry as a result that religious hate speech laws in particular threaten to close down whole areas of edgy but important discourse – especially given that so much religious speech is, nowadays, political speech.

But there is a deeper problem. Hate speech laws are commonly defended by reference to the harm that hate speech does, following the dominant Anglo-Saxon philosophy of criminal law – that its purpose must always be to prevent harm to others.[14] But this is a problematic strategy. As my co-author, Professor Heinze, amongst others has argued,[15] there is an absence of clear empirical evidence indicating the effectiveness of hate speech bans and indeed of the link between hate speech and harms such as discrimination, violence and social exclusion, especially in prosperous, mature, stable democracies.

This is an important point about which it one must be clear: many scholars allow that in politically unstable countries with very strong ethnic tensions, high levels of political violence, and fragile democratic institutions, hate speech may play an important role in inciting terrible violence and thus may need to be restricted. Rawanda is probably the pre-eminent example, but many also believe hate speech played a part in igniting, and then fanning the flames of violence in the terrible breakdown of the former Yugoslavia, and going back further in history, the rise to power of the Nazis. But, says Heinze, this appeal to what he terms 'the catastrophic example' is misguided. What he calls stable, long-standing prosperous democracies are in no danger of collapsing into mass, ethnically-motivated violence or civil war and, with their rich resources in education and culture, have much more effective ways of tackling the harms of hate speech than the blunt and ineffective tool of criminal law.

---

[13] *Hammond v DPP* [2004] EWHC 69 (Admin), applying section 5 of the Public Order Act 1986.
[14] A stance ultimately based on John Stuart Mill's famous 'harm' principle: *On Liberty* (1859).
[15] E. Heinze, "Viewpoint Absolutism and Hate Speech" (2006) 69(4) MLR 543.

On the other hand, when pro-ban scholars, like feminists on pornography use broader arguments about the softer "cultural harm" that pornography does,[16] by constructing the social reality of gender, through pervasive negative and reductive images of women as primarily to be valued for their sexual attractiveness, availability and submission to male desire, the argument becomes hugely over-inclusive. It may plausibly be argued that 'respectable' men's magazines, featuring lingerie-clad models, tabloid newspapers, music videos, Hollywood films and of course, women's magazines themselves, with their often obsessive focus on fashion, the body and sex, do as much if not far more damage to the general perception of women in this way than does violent pornography. The same applies to the treatment of other groups, notably immigrants, in the mass media. As Richard Abel has said: "Were we to take the consequential argument seriously, we would have to accuse mainstream culture rather than scapegoating...targets on its fringe'.[17] Such an argument thus provides little assistance in isolating a particular class of speech as deserving of criminal sanctions.

There is a further difficulty: even if you accept, as I do, that hate speech does cause harm, you then have to confront the question: are hate speech laws effective in reducing that harm? Unfortunately it remains the case that there is virtually no positive empirical evidence that hate speech bans *are* effective in, e.g. minimising harmful messages, enhancing race relations, decreasing discrimination or racially motivated violence. A recent study of Australia's hate speech laws – which was sympathetic to them – found no evidence that they had led to any diminution in racial and other kinds of abuse such as homophobic hate speech.[18] What hate speech laws may do is lead people to express the same basic ideas, but in less crude language – with a greater appearance of sophistication. But this may actually be counterproductive. Given that the use of crude racial epithets is likely to put many readers off, a more apparently reasonable - because more moderate-sounding argument - might actually have wider appeal, thus possibly causing the harmful ideas themselves to gain wider currency.

Hate speech laws can have further, well known, counter-productive effects: prosecutions may turn a hitherto obscure and despised group into 'free speech martyrs', shifting public focus from the false and bigoted nature of their views onto the alleged injustice of their prosecution by criminal law, and allowing them to re-package themselves as romantic revolutionaries against a persecuting orthodoxy. Moreover, prosecution invariably results in vastly greater publicity being

[16] See e.g. "Striking a balance: arguments for the criminal regulation of extreme pornography" Clare McGlynn; Erika Rackley Crim. L.R. (2007), Sep, 677-690
[17] Abel, *Speech and Respect* (1994) at 89.
[18] Copy on file with author.

given to the original message results as it is amplified and repeated in the media, in social media and the blogosphere. Such laws may also miss their intended targets and instead be used against historically oppressed minorities using incautious language to express their anger and alienation: one scholar has pointed out there were almost as many prosecutions of black power activists as white racists under the UK's race hate laws.[19]

III. Some tentative replies on efficacy

There is much in these criticisms but some answers can be given. In particular it can be argued that the line between mature stable democracies and other countries is not as secure as Heinze suggests. I would argue that there can be dangerous drifts even *within* such democracies – perhaps within small geographic or urban pockets of them that suffer from high levels of unemployment and poverty - into intolerance towards certain groups: a recent example concerns the increasing hatred and intolerance expressed towards the Roma people in Italy, which appears to have resulted in violent attacks against them. Moreover, countries can face the quite sudden threat of a collapse in economic and social order – is Greece a stable democracy still? Even the model Scandinavian countries have seen the rapid rise of far-right anti-immigration parties of late. It is frightening how quickly mass employment and poverty can resurrect ghosts of prejudice and hatred we thought we had laid to rest, something the far Right knows only too well.

But even if you can cast doubt on Heinze's confidence about the resistance of democracies to hate speech, you are still faced with the argument that hate speech *laws* are an ineffective or even counter-productive way of dealing with them. However, while I accept much of Heinze's argument here, I think that on this point there is a weakness in the reasoning. Critics like Heinze, looking at hate speech bans *as they are*, conclude that it is not clear they've done any good, while they may have done some harm of the kind they describe. From this (and other grounds) they argue that *all* hate speech bans are undesirable. My response is that this involves a leap in reasoning. To put it crudely, Heinze argue that existing bans are ineffective, *therefore* all bans will be ineffective. I argue that this is not wholly persuasive: one may reasonably infer that narrowly-drafted laws would not cause the same mischief as some of the vague and over-broad laws we see in many European countries. Thus I concede that some actual over-broad or badly conceived bans probably have had counter-productive effects. But I do not accept this means that *all* bans are inevitably ineffective or counter-productive.

[19] Abel (op cit).

IV The necessity for a deontological defence of hate speech bans

However, this response has only got us to the stage of casting doubt on the argument that hate speech bans do no good. And even if you believe that hate speech bans do some good, you would probably have to concede that their effects will only be marginal (it is often pointed out that European hate speech laws have seemed powerless to halt the steady rise of the far Right in many countries in the last ten years). I think you can plausibly argue that hate speech poses at least a *risk* of harm and that hate speech bans *may* help ameliorate this harm. The problem that then arises is that such uncertain gains are not normally considered strong enough grounds to justify restricting a fundamental human right such as freedom of expression. In order to justify such a restriction under international human rights standards – and often constitutional standards too – the state must show that its restriction is *proportionate*:[20] that normally involves showing that the measure is *effective* in some way (an ineffective measure cannot be proportionate, because then the state is restricting a fundamental right for no good reason). And it is very difficult to evidence the claim that hate speech laws are effective in this way.[21] This is why I think you need a solid *deontological* case to demonstrate that such expression should not be regarded as speech at all, because there is no right to utter it; that the state has a moral duty to punish such expression *even if* it is hard to show that such punishment will have clear consequential benefits. That is the case I think we have to make.

Where then, might such a case start? I suggest it might start with the widely-shared insight that hate speech that urges people to regard the target group as non-human, or less than human, is in principle the most dangerous and destructive idea that can be uttered (even if most democracies are can contain the effects of such speech in practice). As one scholar has pointed out, 'an individual cannot enjoy rights in relation to another unless she is recognised as a human.'[22] Therefore the right to be recognised as a person is the most fundamental right of all, for without it, no other rights granted to humans – what we call 'human rights' – will be recognised as inhering in that person. Hence they may then be killed, raped, enslaved, de-citizened. If you read any account of the Second World War and the appalling atrocities accompanying it, whether the Germans against the Jews or Slavs, or the Japanese against the Chinese, you will see, over and over again, that the necessary precondition for inhuman conduct is that the target groups now be

[20] As e.g. required by the Strasbourg Court's interpretation of Article 10(2) ECHR.
[21] This is essentially the argument made by Sumner in relation to the Canadian constitutional context in his essay in Hare and Weinstein, *Extreme Speech and Democracy* (Oxford, OUP, 2009).
[22] S. H. Heyman, *Free Speech and Human Dignity* (Yale UP, 2008) at 000. I am indebted to Heyman for the notion below that hate speech laws should be limited to those that urge the non-recognition of others.

seen as less than human. Richard Rorty, in the opening lines of his Amnesty lecture on human rights, quotes journalist David Reiff, reporting on the conflict in the former Yugoslavia:

> "To the Serbs, the Muslims are no longer human....Serbian murders and rapists do not think of themselves as violating human rights. For they are not doing these things to fellow human beings, but to *Muslims*. They are not being inhuman, but rather are distinguishing between the true humans and the pseudo humans."[23]

Rorty describes this mindset as the belief that "There are animals walking about in humanoid form." Once you're not a human then you're, like the Slavs or Jews to the Nazis, *untermensch* – sub-human, and thus with no rights and without appeal to conscience. In general, in order to get people to do inhuman things to people, you need to get them to stop seeing the victims as human. That way, you try to close off empathy - one of the fundamental bars to inhuman treatment of others.

I suggest that this is *the* dangerous idea – the most dangerous idea in history - and I propose that hate speech laws should have the important but narrow task of opposing it, by insisting that speech may not be used to urge the non-recognition of others or inciting violence against them. Note that I argue this while accepting that, at least in mature stable democracies, most people won't act on this terrible idea. My strategy instead is to argue essentially that there is 'no right' to utter to such speech - that it properly falls outside the human right to free speech because the reasons for valuing free speech in the first place dictate this outcome. Bans on such speech seek to enforce a fundamental premise of the social contract, or of deliberative democracy, namely mutual recognition extended by each to all of a shared citizenship and humanity. I claim therefore that the moral case for such bans exists as part of the same logic that grants us free speech in the first place; in other words, that free speech is as it were, born limited by the conception I describe and that speech that seeks to deny mutual recognition is therefore an abuse of free speech in the true sense of the word.

This is the most important strand in my justification for about speech laws. The second strand draws on the work of the US scholar, S.J. Heyman, to show how this basic duty of mutual recognition has deep roots as a foundational value in our philosophical heritage – particularly in Hegel.[24] The aim is to show that this has long been recognised as fundamental – as important as free speech itself. The third strand draws on the work of Jeremy Waldron,[25] identifying a more

[23] "Human Rights, Rationality and Sentimentality', at http://www.nyu.edu/classes/gmoran/3RORTY.pdf
[24] Heyman, *op cit*.
[25] *The Harm in Hate Speech* (Harvard University Press, 2012).

practical function of hate speech bans as constituting part of the *public assurance* – particularly to minority groups - of their status as possessing a level of basic dignity and equality that all are entitled to. This basic status dignity consists, then, of the same thing that the first two strands are concerned with: public recognition as a human being and a citizen. Hate speech regulation, Waldron argues, can be understood as the protection of 'a certain sort of precious public good': a visible *assurance* offered by society to all of its members that they will not be subject to abuse, defamation, humiliation, discrimination, and violence on grounds of race, religion, and so on. Thus the harm that hate speech laws seek to combat "is the dispelling of the assurance and the dispelling of assurance is the speech act" [26] committed by hate speech.

I concentrate in the remainder of the paper on the first strand.

V. The notion of the duty of mutual recognition as pre-constitutive

The right to free speech is often defended as *pre-constitutive* of the legal order; because we construct the legal order through public discourse or deliberative democracy; this is partly what explains the particular importance afforded to free speech in First Amendment jurisprudence. My argument in response is that the duty of mutual recognition as I've described it is also pre-constitutive. By that I mean that basic equality status is not something that *emerges* from democratic deliberation – what Robert Post calls a 'substantive norm', like decisions about how to provide health care, or what taxes we should have – it is a *formal precondition* for it to take place in the first place. This is most obviously because unless you are recognised as human in the first place you don't get the *human* right to free speech, to participate at all and others will not recognise your contribution as a human point of view. Hence to use speech to attack the basic humanity-status of others is to contradict the premise on which your right rests.

I think the same thing comes out of a non-foundationalist account[27] like that of Jurgen Habermas. Consider his statement that: "the elements of the legal order...presuppose collaboration amongst citizens who recognise one another, in their reciprocally related rights and duties, as free and equal consociates under the law. *This mutual recognition is constitutive for a legal order from which actionable rights are derived*."[28] As Baynes puts it, "[Habermas] emphasises the fact

[26] *Ibid* at 167.
[27] By which I mean that Habermas's vision is a procedural one. "Nothing is given, prior to the citizen's practice of self-determination other than the discourse principle, which is built into the conditions of communicative association in general and the legal medium as such" (*Between Facts and Norms*, 127-28).
[28] *Ibid,* 88.

that rights are not primarily things individuals possess but *relations* that have their basis in a form of mutual recognition...".[29]

So when the state punishes those whose viewpoint effectively is: "these persons are not to be regarded as humans", or "this group should be deported en masse", it is not imposing a substantive norm that ought to be left open to public debate; it is doing what the state must always do: establishing the basic conditions within which debate may take place - that is to be conducted by citizens who mutually recognise each others as free and equal. Hence my core argument is that *speakers urging 'non-recognition' have placed themselves outside the pre-legal community that constitutes their rights*.

VI. How this model provides an answer to the core argument of principle against hate speech bans

So where does this leave the charge by opponents that hate speech laws invade the moral autonomy of the speakers who are punished, thereby failing to treat them with the equal respect that the state must extend to all its citizens - perhaps the key principled objection to such laws? This objection points out that, in a democracy, people have to put up with the fact that they will be out-voted – that laws will be imposed upon them that they don't like and voted against. But it is said, what preserves the legitimacy of the state and the dignity of citizens as moral agents, is that they are always free to express their view upon any matter of public concern: if the state silences them, it becomes, as Dworkin and Heinze have argued, *illegitimate* in relation to them: they are, in a sense, de-citizened. As US 1st Amendment scholar Robert Post puts it: 'Whenever law chooses to enforce cultural norms, as for example, by enforcing norms that distinguish hate speech from normal disagreement, law hegemonically imposes a particular vision of those norms'.[30]

To this I respond, with Heyman, that the recognition of everyone as humans and citizens is not just "one community's norms" or a viewpoint that people should be free to affirm or deny; and nor is it the expression of a particular ideology: it is a norm that is inherent in – indeed constitutive of - *any* 'community', other than one based on tyranny or domination.[31] Indeed I contend, turning Dworkin's legitimacy argument on its head, that the legitimacy of the state *requires* it not to tolerate such speech, precisely because recognition of the other as human is a

[29] K. Baynes, "Democracy and the Rechsstaat: Habermas's Faktisitat und Getung" in White, *Cambridge Companion to Habermas* at 209.

[30] See e.g. Post's essay in Hare and Weinstein, *op cit* and his "Racist Speech, Democracy and the First Amendment" 32 Wm & Mary L. Rev. 267-25 (1991).

[31] Heyman, in Hare and Weinstein at 173.

pre-condition for any kind of non-violent, consensual deliberation and indeed is prior to any system of human rights that might emerge from that deliberation. In short then if speech urging 'non-recognition' is banned, what is then happening is not that a specific, substantive set of norms is being allowed to shape the public sphere, to the detriment/exclusion of groups who don't share it, but that a basic precondition for democratic deliberation is necessarily shaping the parameters of public discourse.

The main response to this kind of argument – the tack that Dworkin, Heinze and Edwin Baker[32] take – is to say that I am simply confusing the duties of the state and the individual. The *state* of course must respect the equal dignity of its citizens; but individuals must be free to dissent. Heinze argues that only the state can actually *deny* equal citizenship, through e.g. enacting discriminatory laws; other citizens voicing hateful speech cannot *deny* such citizenship, they can merely *disdain* it.

There are two rejoinders to this important counter-argument. One I take from Waldron – what I think is his most important insight in this debate: that the implicit assurance of the equal basic dignity of all citizens is a public good, but not one that can be supplied by the state acting alone. Unlike, say, electricity, which can be provided by one single body, this public assurance is something that 'arises out of what thousands or millions of ordinary citizens do singly and together. It is, as Rawls puts it, a product of "citizens' joint activity in mutual dependence on the appropriate actions being taken by others."[33] So this a duty – a negative one, not to undermine this public assurance through publicly uttered hate speech – is something that *must* be laid on citizens, because the state cannot achieve it alone.

The second rejoinder simply refers back to the previous argument: if mutual recognition really is a foundational value –an essential condition for deliberative democracy - then we have the same conclusion: logically, the state *cannot* supply it alone: for we are talking not about state-citizen recognition, but inter-citizen recognition. The state must therefore do what it can to ensure that citizens do not themselves actively undermine such recognition. In doing so it is not interfering with merely private act, or engaging in thought control, but rather doing what the state must do: upholding a basic but very precious public good. This is a public good that at the very least, citizens must be obliged not to actively undermine. While the state has positive duties to uphold it, citizens have at least a negative duty not to undermine it.

[32] See his essay in Hare and Weinstein, *op cit*.
[33] Waldron, *op cit* at 93 (the reference to Rawls is to *Political Liberalism* at 204).

VII The model of hate speech that follows

So what model of hate speech does this conception give rise to and what does it cover? In explaining how this differs conceptually than many existing hate speech laws I find very helpful Stephen Darwall's distinction between 'recognition respect' and 'appraisal respect'.[34] The former is fundamental to the dignity of persons and invariant in the face of differential merit. The latter is respect that is earned and assessed by conduct, beliefs and character.[35] I believe that many hate speech laws get in a muddle, and try to do more than they should, by trying to protect both kinds of respect: so a core argument of my conception is that they should be concerned only with "recognition respect". They should enforce the narrow but fundamental duty to recognise others as humans and citizens and concomitantly the duty not to call for violence against such groups. This conception, because it does not in any way seek to close down debate about which forms of life are most valuable, or indeed harmful or contemptible, is intrinsically narrow: it does *not* include within its scope speech that criticises, insults or ridicules – however unfairly - conduct or belief, particularly in the religious, but also in the sexual sphere. The law should *not* then try to make you give equal respect to all religions or even adherents of different religions.[36] Indeed I would argue that adherents to the Church of Scientology, or the fundamentalist Christian right may reasonably be thought to require 'severe ridicule'.[37] We have to not only oppose Islamic fundamentalism but laugh at it and its adherents as well and the law should not even come close to trying to stop us from doing so.

I conclude by trying to draw some lines to indicate what would be covered by my model, and what would not.

Hate speech laws *should* cover three things:

> **First**, speech that, explicitly or implicitly denies the equal humanity of the target group for example, by referring to them as germs, cockroaches or vermin. Note however that no obviously insulting terms of this sort would *have* to be used. A reasoned argument to the effect that black people were less than human would be caught. It would not be necessary to use insulting racial epithets.

[34] 'Two Kinds of Respect' *Ethics* 88 (1977).
[35] I am grateful to Waldron for drawing my attention to this important work: *op cit* at 87.
[36] I have in mind here the (unsuccessful) prosecution of Michel Houellebecq for remarking that Islam was 'the stupidest of all religions'.
[37] One of the definitional phrases used by the state hate speech law of the state of Victoria, Australia.

Second, it would catch speech that advocates violence against an identifiable group. So I approve the recent conviction in the UK of some Muslim men for handing out leaflets calling for the execution of gay people.[38]

Third, it would cover speech that does not recognise members of a given group as citizens – this means in practice calling for such groups to be wholly stripped of their ordinary constitutional rights – as in calls for mass deportation, enslavement, confiscation of property and so on. Thus **t**he statement prosecuted in the English case of *Norwood* v *DPP*:[39] "Muslims out of Britain", would be covered, if in the context it was clear that the call was for mass deportation.[40]

What then would *not* be included in my model?

First, insults as such, unless the insult means that the speaker falls within one of the above categories. (Recall that face-to-face insults or targeted insults, e.g. in the form of phone calls or emails, are catered for by other laws).

Second, in relation to the sphere of religious hate speech, attacks on doctrines, beliefs and practices *of any sort* would not fall within the definition. Neither would religiously derived statements to the effect that certain sexual practices such as gay sex are sinful or immoral be covered. Take, for example, statements sometimes made by the religious Right in the US that 'Fags burn in hell'. While hate speech laws often cover 'threatening words', in my view, threatening divine retribution cannot be counted as a threat. (To be legally recognised by a secular state, 'threats' have to be of something happening in this life and not of action by a divine being). On the other hand, the mere fact that the statement is connected to religious belief does *not* of itself afford a defence: so when certain African Christian clergy refer to gay people as cockroaches and call for their execution, that *would* be caught by my model; as would Islamist preachers calling for the killing of unbelievers.

Third, calls to restrict immigration, even of a particular group, should *not* be covered. A rightwing Dutch MP was prosecuted for racial insult in the 1980s for saying "Full is Full"; such statements would plainly not fall under my model.

---

[38] *Ali, Javed and Ahmed* (unrep.) 10 Feb, 2012.
[39] (2003) WL 21491815
[40] I have not made up my mind as to whether so-called 'group defamation' – the making of factually false claims about a group that are likely to bring it into hatred or contempt, should be covered by hate speech law.