



Jurisprudence

An International Journal of Legal and Political Thought

ISSN: 2040-3313 (Print) 2040-3321 (Online) Journal homepage: <https://www.tandfonline.com/loi/rjpn20>

Is the 'hate' in hate speech the 'hate' in hate crime? Waldron and Dworkin on political legitimacy

Rebecca Ruth Gould

To cite this article: Rebecca Ruth Gould (2019): Is the 'hate' in hate speech the 'hate' in hate crime? Waldron and Dworkin on political legitimacy, *Jurisprudence*

To link to this article: <https://doi.org/10.1080/20403313.2018.1552468>



Published online: 22 Jan 2019.



Submit your article to this journal [↗](#)



View Crossmark data [↗](#)



Is the 'hate' in hate speech the 'hate' in hate crime? Waldron and Dworkin on political legitimacy

Rebecca Ruth Gould 

College of Arts & Law, University of Birmingham, Birmingham, UK

ABSTRACT

Among the most persuasive arguments against hate speech bans was made by Ronald Dworkin, who warned of the threat to political legitimacy posed by laws that deny those subject to them adequate opportunity for dissent. In his influential defence of hate speech bans, Jeremy Waldron addresses these objections. Dworkin's concern with political legitimacy is misplaced, he argues, given the provision speech bans make for substituting permissible modes of expression for impermissible ones. I argue that this defence of speech bans misidentifies the 'hate' in hate speech with the 'hate' in hate crime. In contesting Dworkin, Waldron fails to contend with the necessarily entangled criminalisation of manner and viewpoint entailed in hate speech bans. By failing to grapple with the way in which every linguistic sign is constituted by both manner and viewpoint, Waldron sidesteps the ways in which hate speech bans undermine political legitimacy within liberal democracies.

KEYWORDS

Hate speech; censorship; democracy; political legitimacy; Ronald Dworkin; Jeremy Waldron; IHRA definition; antisemitism; Israel/Palestine

In his influential defence of hate speech bans, *The Harm in Hate Speech* (2012), legal philosopher Jeremy Waldron addresses the most formidable arguments that to date have been made in support of their abolition. Among the most compelling of the exchanges contained in that book is with his teacher Ronald Dworkin, many of whose positions Waldron follows and develops, but from whom he diverges with respect to his position concerning hate speech regulation.¹ The Dworkin-Waldron debate was rekindled five years later in an exchange between Jeremy Weinstein and Waldron, in which Weinstein develops Dworkin's position, and Waldron restates (and, in some cases, reformulates) the core argument of his 2012 book.²

While Dworkin maintains a prominent position in the pantheon of legal philosophers, in more broadly sociological terms, Waldron's support for the regulation of hate speech is the mainstream view at present, both within the context of liberal European legal theory as well as among legal scholars who position themselves to the left of liberal theory.³ In light

CONTACT Rebecca Ruth Gould  r.r.gould@bham.ac.uk

¹For Waldron's appreciation of Dworkin's legacy, see Jeremy Waldron, 'Jurisprudence for Hedgehogs' (2013) New York University Public Law & Legal Theory Research Paper Series Working Paper No. 13-45.

²See James Weinstein, 'Hate Speech Bans, Democracy, and Political Legitimacy' (2017) 32 *Constitutional Commentary* 527 and Jeremy Waldron, 'The Conditions of Legitimacy: A Response to James Weinstein' (2017) 32 *Constitutional Commentary* 697.

³A measure of the positive reception accorded *The Harm in Hate Speech* can be seen in Julian Rivers, Tariq Modood, Simon Thompson, and Karen Zivi, 'Understanding and Regulating Hate Speech: A Symposium on Jeremy Waldron's *The Harm in*

of the consequences that Waldron's defence of speech bans has for broader understandings of the relationship between the regulation of speech and democratic legitimacy, and ultimately for the meaning of democracy, it is worth dwelling in detail on the grounds of his disagreement with Dworkin, and considering to extent to which his defence answers Dworkin's legitimacy-based critique.⁴

In reviewing various arguments linking free speech to democracy (Meiklejohn, Baker), Waldron notes that, in failing to go beyond a general concern for the democratic process, opponents of hate speech bans address the issue in such generalised terms that they fail to identify the specific challenge to democratic legitimacy posed by hate speech. Waldron then turns to Dworkin's argument against hate speech bans, which he takes to be both more specific and more challenging than those that have been advanced before. In Waldron's characterisation, Dworkin argues that 'free and unrestricted public discourse is a *sine qua non* for political legitimacy in a democracy, not just for the quality of democratic engagement'.⁵ Waldron then sharpens the case he seeks to answer: '*the legitimacy of certain specific legal provisions may be imperilled by the enactment and enforcement of hate speech laws*' (174, emphasis in original). He cites from *Extreme Speech and Democracy* (2009), a landmark volume edited by Ivan Hare and James Weinstein, where Dworkin argues:

Democracy requires ... that each citizen have not just a vote but a voice: a majority decision is not fair unless everyone has had a fair opportunity to express his or her attitudes or opinions or fears or tastes or presuppositions or prejudices or ideals ... to confirm his or her standing as a responsible agent in, rather than a passive victim of, collective action.⁶

As suggested in this quote (if in underdeveloped form), Dworkin believes that hate speech bans undermine democratic legitimacy to the extent that they deprive the citizen of a voice in the political process. They deny to individuals who hold views targeted by hate speech bans the opportunity to speak without fear of criminal sanction. Dworkin had made this point already in 1994, and again in 2006, when he argued in *Index on Censorship* that 'it is illegitimate for governments to impose a collective or official decision on dissenting individuals, using the coercive powers of the state, unless that decision has been taken in a manner that respects each individual's status as a free and equal member of the community'.⁷ It is through the process of voicing one's views (and not merely voting), Dworkin argues, that the citizen acquires a sense of political responsibility that also entails a duty to obey the law.

On Dworkin's account, a state that denies to the citizen a forum for expressing her prejudice, her hate, and even her racism without fear of criminal sanction also denies to

Hate Speech' (2014) 13 *Contemporary Political Theory* 88 (Dworkin's position goes unmentioned here). For the resonance between Waldron's views and European jurisprudence, see the fact sheet 'Hate Speech' produced by the European Court of Human Rights <www.echr.coe.int/Documents/FS_Hate_speech_ENG.pdf> accessed January 2019.

⁴Although I draw here primarily from Waldron's discussion of Dworkin in *The Harm in Hate Speech*, I also rely on another iteration of this argument, published the same year as Waldron's book, and included in *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, Michael Herz and Peter Molnar (eds.) (Cambridge University Press 2012); Jeremy Waldron, 'Hate Speech and Political Legitimacy' 329–40, with a response by Ronald Dworkin, 'Reply to Jeremy Waldron' 341–44.

⁵Jeremy Waldron, *The Harm in Hate Speech* (Harvard University Press 2012) 174.

⁶Ronald Dworkin, 'Preface' in Ivan Hare and James Weinstein (eds), *Extreme Speech and Democracy* (Oxford University Press 2009) v–ix. This work adapts some of the text in 'A New Map of Censorship' (n 7).

⁷Ronald Dworkin, 'A New Map of Censorship' (2006) 35 *Index on Censorship* 131. This article is an abbreviated version of an article published under the same title also in *Index on Censorship* 1/2 (1994) 9–15.

her the capacity to act as a responsible agent in the democratic political process. Given that political legitimacy is a matter of degree rather than of kind, the illegitimacy entailed in hate speech bans does not of itself license a violation of the legal order. Yet hate speech bans inevitably leave us, in Dworkin's words, with 'something morally to regret'; they generate 'a deficit in legitimacy' whenever and wherever they are implemented.⁸ This deficit makes them inconsistent with democracy.

While a state does not become illegitimate through hate speech bans alone, any legislation that bars citizens from protesting the laws they are expected to obey generates a legitimacy deficit, and may eventually precipitate a legitimacy crisis. Further, the type of protest that democracies must permit as a condition for their legitimacy is not containable within the bounds of 'civil discourse', as the term is widely understood. Protest permissible within a democracy includes anger, hate, and vituperation, each of which has a political role within public discourse, yet each of which is regularly targeted by speech bans.⁹ For Dworkin, any attempt to suppress uncivil forms of expression that do not directly threaten citizens' security compromises the legitimacy of the democratic state.

On this Dworkinian account, hate speech bans pre-empt the democratic process while failing to offer citizens, as egalitarian legal regimes must, the opportunity to claim their political agency by contesting the laws by which they are governed. Speech bans' compromise to political legitimacy transpires across two domains in Dworkin's account, as reframed by Waldron.¹⁰ The first compromise is procedural; it refers to the processes that ratify legislation and which give citizens the opportunity to contest the terms of legal provisions, generally through an elected representative. As noted by Waldron, Dworkin regarded compromises on this type of legitimacy as consistent with democratic outcomes, in his defence of judicial review.¹¹ While procedural legitimacy is desirable in a democracy, it is neither a necessary nor a sufficient condition for this political form.

The second kind of legitimacy is symbolic, and is encapsulated in what Waldron calls 'the basis of a state's right to govern'.¹² Although Waldron does not discuss this type of legitimacy in detail, it reaches well beyond proceduralism, and is primarily secured symbolically, including by the perception the citizen has of being included within the social order. Proceduralism is useful in securing symbolic legitimacy, but it is not a sufficient condition for it. Symbolic legitimacy requires unlimited scope for speech acts that are not subject to viewpoint-selective censorship. Speech bans undermine democratic legitimacy procedurally and symbolically, and consequently weaken the citizen's obligation to obey the laws promulgated by such states.

While procedural illegitimacy can be addressed through compensatory measures (notably judicial review), the damage to symbolic legitimacy that inheres in speech bans

⁸Dworkin, quoting from his own email to Jeremy Waldron, in 'Reply to Jeremy Waldron' 341.

⁹On the political value of anger, see Amia Srinivasan, 'The Aptness of Anger' (2018) 26 *Journal of Political Philosophy* 121, and Maxime Lepoutre, 'Rage Inside the Machine: Defending the Place of Anger in Democratic Speech' (2018) 17 *Politics, Philosophy, & Economics* 398.

¹⁰While the distinctions between different types of legitimacy discussed here are inspired by both Dworkin and Waldron, they are my own.

¹¹The reference is to Ronald Dworkin, *Freedom's Law: The Moral Reading of the American Constitution* (Oxford University Press 1996) 15–35. Waldron (in 'Hate Speech' 698) presents Dworkin's free speech legitimacy thesis as evidence for an evolution in his views in relation to legitimacy; however, as per n 7, Dworkin had already begun to articulate his perception of the link between free speech and legitimacy in 1994, two years prior to his defense of judicial review's compromise with procedural legitimacy in *Freedom's Law*.

¹²Waldron (n 3) 698.

is impossible to remedy. Speech bans always and necessarily compromise the symbolic dimension of legitimacy. This compromise is more problematic for democracy than is procedural illegitimacy, which is always subject to correction, as part of perpetual negotiations between the citizen and the state. Far from securing democracy for its citizens, speech bans undermine the foundations of democratic legitimacy for everyone, and imperil the state's ability to legitimately command obedience to its laws.¹³

Waldron deserves credit for drawing attention to Dworkin's legitimacy thesis, the importance of which had prior to him been underestimated. As he wrote five years after the publication of *The Harm in Hate Speech*, 'I like to think I have contributed something to this increase in rigor [within the hate speech debate] by subjecting Dworkin's version of the legitimacy argument to sustained exposition and critique.'¹⁴ Addressing the revival of the debate in his exchange with Weinstein, Waldron notes that 'we have a version of [Dworkin's] argument ... presented in a sustained and rigorous way, and the debate about free speech and hate speech is the better for it'.¹⁵ Through his detailed and informed engagement with Dworkin's tersely expressed views, Waldron has at once critiqued Dworkin while also helping to reveal the relevance of his thinking to the contemporary debate around hate speech and censorship.

Waldron follows Dworkin closely up to the point of their divergence, distinguishing between downstream laws targeting violent acts, that would be classified as hate crimes, and upstream laws, targeting hate speech. Waldron further clarifies that whereas downstream laws are 'enacted by the political process', upstream laws 'affect the political process'.¹⁶ While both Dworkin and Waldron insist that the state must protect citizens 'against unfairness and inequality in employment or education or housing or the criminal process', only Dworkin advises against intervention 'further upstream, by forbidding any expression of the attitudes or prejudices that we think nourish such unfairness or inequality'.¹⁷ In adopting Dworkin's distinction between upstream and downstream legislation, Waldron implicitly recognises that hate speech and hate crime laws are not identical. Indeed, in *The Harm in Hate Speech*, he insists on this distinction:

Though the two ideas – hate speech and hate crimes – do have a distant connection, they really raise quite different issues in our thinking about law. The idea of hate crimes ... definitely does focus on motivation: it treats the harboring of certain motivations in regard to unlawful acts like assault or murder as a distinct element of crime or as an aggravating factor. But in most hate speech legislation, hatred is relevant not as the motivation of certain actions, but as a possible effect of certain forms of speech. (35)

Yet, because he fails to think through the upstream/downstream distinction in terms of the difference made by language, hate speech merges with hate crime in interesting and problematic ways in Waldron's effort to refute Dworkin.

¹³A similar argument has been advanced more recently by Eric Heinze in *Hate Speech and Democratic Citizenship* (Oxford University Press 2016). Whereas Dworkin articulates his case within a liberal rights framework, Heinze proposes instead that we view free speech 'not only as an *individual right*, but as an *essential attribute of democratic citizenship*' (4, emphasis in original).

¹⁴Waldron (n 3), 698.

¹⁵ibid 697–98.

¹⁶ibid 704. Of course, it is possible to question whether the distinction between being enacted within a political process and affecting a political process adequately captures Dworkin's idea, and to wonder why these two aspects should be seen as mutually exclusive.

¹⁷Dworkin (n 7) 132.

I show here that Waldron's neglect of the constitutive difference between the 'hate' in hate speech and the 'hate' in hate crime undermines his refutation of Dworkin's legitimacy thesis. Further, collapsing this distinction enables Waldron to conclude erroneously that 'banning hate speech ... has no greater effect on political legitimacy than banning fighting words or these other acknowledged exceptions to the free-speech principle' (183). Before proceeding with this argument, it will help to clarify how the concepts that structure this inquiry are understood here. Although they are often discussed together and the distinction between them is sometimes blurred, hate speech and hate crime are constitutively distinct.

Defining hate crime

The Crown Prosecution Service defines a hate crime as 'any criminal offence which is perceived by the victim or any other person, to be motivated by hostility or prejudice'.¹⁸ The opening clause in this definition – 'any criminal offence' – is significant; hate crimes are limited to acts that can formally be classed as criminal offences. An act must first be established as a crime before it can properly be considered as a hate crime. Equally, no act that is not a criminal offence is eligible for membership in the category of hate crime. Racial prejudice functions in this context as evidence for *mens rea* (guilty mind), which, when it occurs in union with *actus reus* (guilty act), is constitutive of a crime.¹⁹ Criminality requires both intent and the commission of an act; neither can be sufficient in isolation from the other. By definition, all hate crimes are criminal acts, but not all criminal acts are hate crimes.

The role of hate in constituting hate speech (as construed by hate speech bans, including the UK's The Racial and Religious Hatred Act of 2006) is radically distinct from the role of hate in constituting a hate crime. While both the courts and existing legislation in any given jurisdiction may conflate these two, there is no reason for legal theory to perpetuate this error.²⁰ Whereas a hate crime must first be classified as a crime before it can be treated conceptually as a hate crime, the 'hate' in hate speech is not something that is later added in order to constitute it as a crime. Rather, in the case of hate speech, hate – combined with its vituperative manner of expression – is *constitutive* of the status of a speech act as a crime. While hate is intrinsic to the (perceived) criminality of hate speech, it is merely an aggravating factor in determining the severity of a hate crime. Further, as Eric Heinze as notes, in the context of criminal adjudication, the 'tort of intentional infliction of emotional injury' has long been available to criminal prosecutors as additional evidence of criminal intent.²¹ Any refusal to take evidence of racial hate into account in criminal adjudication would be both anomalous and discriminatory. The burden of the argument lies with those who oppose such an application of hate crime legislation, not with those who support it.

¹⁸'What is hate crime?' Metropolitan Police <www.met.police.uk/advice/advice-and-information/hco/hate-crime/what-is-hate-crime/> (accessed 12 October 2018).

¹⁹George P. Fletcher, *The Grammar of Criminal Law: American, Comparative, and International* (Oxford University Press 2007), vol. 1, 95.

²⁰Heinze correctly notes that 'States maintaining hate speech bans often classify hate speech as one form of hate crime, suggesting that hateful expression itself already constitutes a harmful act without having to attach to any further material harm' (Heinze (n 13) 19). However, far from limiting the plausibility of the hate speech/hate crime distinction, its blurring in contemporary jurisprudence makes distinguishing between the two all the more urgent as a matter of public policy.

²¹Eric Heinze, 'Viewpoint Absolutism and Hate Speech' (2006) 69 *Modern Law Review* 576. Also see Laurence Tribe, *American Constitutional Law* (2nd edn, Foundation Press 1988) 838 n 16 (cited in Heinze, *ibid*).

While Waldron recognises that hate speech bans are necessarily content-restrictive, he underestimates the extent to which they are viewpoint selective. ‘The kind of restriction we have in mind,’ he writes, ‘operates and is envisaged explicitly as a limitation based on content’ (150). In his later treatment of the topic, Waldron even more explicitly distances speech bans from viewpoint-based censorship. ‘Hate speech restrictions,’ he argues, ‘are not based on viewpoint *per se*, but on the manner of their expression and the effect they are intended to have on social peace’.²² Acknowledging that his proposal to ban hate speech ‘flies directly in the face of one of the pillars of American free-speech doctrine’ (150), Waldron makes explicit his preference for a legal system that permits viewpoint discrimination.

The constitutive link between viewpoint and manner

Waldron does not deny the constitutive role that hate plays in hate speech bans’ criminalisation of speech when he is in the process of defending them. Yet, when he undertakes to refute Dworkin’s legitimacy thesis, he glosses over the way in which language is constituted as much by its meaning as it is by its modality of expression. As millennia of literary and linguistic theory have recognised, the link between viewpoint and manner of expression in every speech acts is *constitutive* in determining a mode of expression; neither can be conceptualised in isolation from the other.

The constitutive link between viewpoint and manner of expression can be illustrated through a highly contested speech ban that was adopted by the UK government in 2016 without being formally incorporated in its legal system: the working definition of antisemitism proposed by the International Holocaust Remembrance Alliance.²³ One of the most controversial provisions contained in this definition is its stipulation that ‘claiming that the existence of a State of Israel is a racist endeavor’ may serve as an example of antisemitic intent. This definition has yet to be formally incorporated into legislation, yet it heavily shapes para-legal and quasi-legal discourse pertaining to contemporary antisemitism, and many of those who advocate the definition support its incorporation into law. Francis Kalifat, leader of the Conseil Représentatif des Institutions juives de France, has for example stated that ‘Our hope is to see this definition integrated into French law.’²⁴

Supporters of the IHRA definition address the charge that it chills free speech by claiming that it is possible to refer to certain Israeli policies as racist without manifesting antisemitic intent. In order to avoid being captured by the definition, an Israel-critical speech act on this reading would need to be appropriately modulated in a way that avoids causing offense to the audience.²⁵ Proponents of the definition claim that, while

²²Waldron, ‘Hate Speech,’ 713.

²³The definition (available at <www.holocaustremembrance.com/sites/default/files/press_release_document_antisemitism.pdf>), has been discussed extensively in Rebecca Gould, ‘Legal Form and Legal Legitimacy: The IHRA Definition of Antisemitism as a Case Study in Censored Speech’ *Law, Culture and the Humanities* (Online First: <https://doi.org/10.1177/1743872118780660>).

²⁴Shirli Sitbon, ‘Despite a Call from French Jews, Macron and Le Pen Have Not Adopted the IHRA Antisemitism Definition’ *The Jewish Chronicle* <www.thejc.com/news/world/despite-call-from-france-jews-macron-le-pen-parties-have-not-adopted-ihra-antisemitism-definition-1.470306> accessed 27 September 2018.

²⁵I do not claim that Waldron, who carefully distinguishes between the causing of offense and the infliction of harm (see *The Harm in Hate Speech*, 129–20), would formulate the requirement in this way. I do however insist that, outside the realm of normative legal theory, this is how such speech bans are implemented. Further, I claim that it is *intrinsic* to the nature of speech bans that they will be interpreted and applied in this way.

it may be permissible to criticise specific Israeli laws as racist, the state's commitment to eradicating racism forbids the blanket denomination of Israel as an apartheid state. From the perspective of those who support the IHRA definition, the deliberately provocative flyers that appeared in various London boroughs following the adoption of the definition by the Labour Party in September 2018 (Figure 1) ought to be situated beyond the boundaries of permissible speech.²⁶ The incident raises the question: is permitting the propositional content of the notion that 'a State of Israel is a racist endeavour' – minus any vituperative manner of expression – adequate to ensuring freedom of expression in this area? Waldron assures us that well-drafted hate speech bans achieve such feats of equivalence.

Were the colouring on the 'racist endeavour' sign muted, the white background less stark, the font dimmer, or its placement less prominent, this sign would have been less likely to have been targeted for criminal sanction by the London authorities. Perhaps too a less categorical statement, such as 'Israeli policies may be racist,' would have mitigated the offense caused. But on what reasonable account of representation could we accept that a reconfiguration of the form through which the statement 'Israel is a racist endeavour' was expressed could ever equal its propositional content? No surgical division between message and manner can be sustained within any form of aesthetic representation. The tone of voice we use, the rudeness of our language, and the intensity of our rhetoric are all part of what constitutes our speech acts. These elements are all equally intrinsic to the ways in which we express ourselves. Any adjustment of the manner of expression to accommodate what Waldron refers to as the 'social peace' will necessarily mute our meaning and alter its propositional content. This is not to say that such adjustments may not be advisable in certain contexts, or that we might not wish to alert our friends, family, colleagues, strangers, and ourselves to the harm that may be generated through the manner in which our messages are expressed. At issue in the case of hate speech bans is not whether certain modes of expression are advisable, however, but whether there is a case for their legally mandated and coercively enforced prohibition by the state.

All viewpoints are intrinsically linked to the manner of their expression. In the case of the Israel-critical speech targeted by the IHRA definition, the viewpoint that proponents of the definition seek to censor may be merely offensive or, maximally, antisemitic; in neither case is this an argument for its censorship. Most philosophers of language recognise the impossibility of imposing a surgical separation between the manner of expression and its propositional content. Hate speech ban advocates, however, appear to have ignored this constitutive aspect of the speech act. For all of these reasons, I concur with Heinze that Waldron's 'form-substance distinction ... derails' his defense of hate speech bans.²⁷

²⁶The IHRA definition does not fit the classical model of a hate speech ban for a range of reasons, most notably its self-described 'legally non-binding' status. This example is however instructive because the lack of clarity around its legal status supports the arguments of those who insist that indeterminacy is intrinsic to the very idea of a speech ban. Equally, even though the UK government is not in a position to incorporate the definition into legislation, the anonymous individuals who posted these posters are threatened with criminal penalties if caught (technically because they were posted without a license on rather than due to the viewpoint they express; see n 28).

²⁷Eric Heinze, 'Taking Legitimacy Seriously: A Return to Deontology' (2017) 32 *Constitutional Commentary* 632.



Figure 1. 'Israel is a Racist Endeavour' sign at a London bus stop (5 September 2018). I thank Nael Alqtati and Malaka Mohammed Shwaiikh for help in obtaining this image.

Legitimacy and the question of settlement

Engaging in racist speech may legitimately ostracise and isolate the racist. The moral wrongness and ethically indefensibility (as well as sheer stupidity) of hateful speech acts may be legitimate grounds for social exclusion as well as for non-employment and social shunning, but, within a democratic society, no viewpoint, and no manner of expressing that viewpoint, is of itself legitimate grounds for legal sanction (absent a recognised exception to the free speech principle, such as incitement to violence). Waldron cautiously dissents from this view in his initial critique of Dworkin by tying legitimacy to the extent to which a contentious issue may reasonably be regarded as settled, in the sense of no longer up for debate, within a given society. Waldron writes:

If the proposal were to ban people from expressing contemptible views about welfare recipients or democratic socialists, then I think there would be a case to be made along the lines of ... Dworkin's argument ... such suppression would put in question the legitimacy of our pursuit of policies based on premises that people were being fined or put in jail for (vituperatively) denying. But we are ... talking about the fundamentals of justice, not the contestable elements. By the fundamentals of justice, I mean things like elementary racial equality, the basic equality of the sexes, the dignity of the human person, freedom from violence and intimidation, and the like.²⁸

In *The Harm in Hate Speech*, Waldron notes that he is no longer sure that he wants to commit himself to the position that when an issue is no longer debated in society, 'we should be less solicitous of political legitimacy when we decide how to deal legislatively with the harm inflicted on the dignity of minority members' by hate speech that supports

²⁸Waldron (n 4) 336.

that discarded view (196).²⁹ Yet, while he expresses caution, Waldron does not repudiate this view. His hesitation notwithstanding, Waldron's impulse to link political legitimacy to the question of settlement is perhaps even more instructive than he imagined.³⁰ Where Waldron errs is in his implication that the more settled a point of view within society and within the law, the less compromise is made with democratic legitimacy when the views to which it is opposed are legally regulated.

To the extent that we can conceptualise legitimacy as a continuum rather than a dichotomy, I draw the opposite conclusions from Waldron's formulation. Waldron understands settlement in two senses: (1) as representing a view 'that people rely on comprehensively and diffusely in almost every aspect of their dealings with others'³¹; (2) as representing the premises of modern social and legal organisation. In the case of the racist speech targeted by hate speech bans, Waldron states, 'If one cannot exact respect for one's basic status as a rights-bearing individual, then almost everything is thrown into question.'³² While I find Waldron's formulation of the relationship between settlement and legitimacy compelling, it seems to me to better support precisely the *opposite* claim: the *less* settled a given point of view, and the more it is subject to contestation, both in terms of social relations and cultural capital and in terms of social and legal organisation, the greater is the compromise to political legitimacy involved in banning that view. We should be wary of censoring unpopular and unsettled views, even more than popular ones.

When a debate is truly settled, it follows that harms cannot accrue from permitting the articulation of the defeated point of view. The case against banning hate speech is linked in part to the relatively settled nature of the debate around racial equality. In a state not proactively committed to 'elementary racial equality, the basic equality of the sexes, [and] the dignity of the human person' the citizen's prerogative of free expression would have limited value. It is only within a democracy – by definition a state committed to racial, sexual, and social equality – that free speech can be seen as foundational to its legitimacy. The free speech mandate, including its contribution to political legitimacy, assumes that the discussion concerns a democratic polity. A state such as Nazi Germany or apartheid South Africa that could not guarantee citizen's basic rights and which refused to implement a policy of equality for all citizens, would be unable to sustain and nurture the democratic value of free speech (including racist speech). Hence, the argument against hate speech bans presupposes a state that is committed to eliminating discrimination. In other words, it assumes a democracy.

Dignity and hate speech

Liberal legal theorists such as C. E. Baker closely link the capacity for expression and with human freedom, and hence with personhood.³³ The formative role of expression in

²⁹Waldron appears to be referencing his earlier article 'Dignity and Defamation' (2010) 123 *Harvard Law Review* 1596. His position concerning the 'settled' status of the debate around equality is criticized in Stephen Holmes, 'Waldron, Machiavelli, and Hate Speech' in *The Content and Context of Hate Speech* 350.

³⁰For further discussion in the context of free speech and racist speech, see David Kretzmer, 'Freedom of Speech and Racism' (1987) 445 *Cardozo Law Review* 445.

³¹Waldron (n 4) 336.

³²*ibid* 336.

³³C. E. Baker, 'Hate Speech' in *The Content and Context of Hate Speech*, 37–80.

shaping personhood means that infringements by the state on the citizen's capacity for expression are tantamount to a denial of rights comparable to the more widely registered denial of dignity entailed in hate speech.³⁴ As Weinstein notes in his reformulation of Dworkin's legitimacy thesis: 'To the extent that such censorship prevents people from expressing what they believe is best for society, it is insulting; in so far the speech restriction impairs their ability to promote or protect their own self-interest, it is also fundamentally unfair.'³⁵

The duality of dignitarian arguments is evident to anyone following the free speech debate. Given a non-neutral state committed to social, racial, and economic justice (all of which are also conditions for democratic legitimacy), the question becomes whether we prefer for the state to take the side of the speaker of hate speech or its target. With speech ban proponents, I prefer for the state to side with the latter, and am not troubled by the possibility that the racist may have to contend with a state that promulgates policies hostile to his point of view. What troubles me is that the state may use 'equality and diversity' as an excuse to arrogate to itself the power to dictate, not just its own point of view, but also the point of view of the speaker of racist speech. Far from being a mere hypothetical, any attentive observer of politics within liberal democracies will be deeply familiar with the state's propensity for engaging in this type of dissimulation.

While the effects of denying racists the right to speak on the one hand and permitting the infliction of hateful words on another are dissimilar, the infringement on the citizen's speech prerogative bears comparison with the indignity entailed in racist speech. The *constitutive* role of hate in hate speech accounts in part for its harms but also lies at the core of its link to political democratic legitimacy. In denying to someone the right to speak in a vituperative manner that would not be treated as criminal (evidence of *mens rea*) in the absence of a racist viewpoint, we also deny to that person the capacity to express their views, and to have them subjected to public debate, without fear of legal sanction. While most European politicians and jurists, including within the Council of Europe and the EU, do not accept an inalienable right to express racist views, it is inaccurate to claim that viewpoint-selective censorship is not entailed in the denial of this right.

Waldron's assurance that hate speech legislation offers to every citizen denied the right to vituperative expression another 'roughly equivalent expression that will not incur legal sanctions' (183) relies on a flawed understanding of the role of medium and manner in constituting expression. Just as a Beethoven symphony performed on a harmonica will never be a 'rough equivalent' to the same symphony performed by the Berlin Philharmonic, racist speech expressed in a vituperative manner cannot be rendered in a more innocuous medium without having its intrinsic meaning – including its propositional content – transformed. Translators have long recognised the untranslatability of all expressive statements.³⁶ Hate speech ban advocates who assume that it is possible to surgically separate form from content, and who ignore the ways in which the two are constitutive of each other, could learn a great deal from approaches that have long been inter-nalised within translation studies.

³⁴Along with *The Harm in Hate Speech* (105–43), another persuasive case for hate speech as a denial of dignity is made in Stephen Heyman, *Free Speech and Human Dignity* (Yale University Press 2008).

³⁵Weinstein (n 2) 540.

³⁶For the history of this position within translation theory, see Rebecca Ruth Gould, 'Inimitability versus Translatability: The Structure of Literary Meaning in Arabo-Persian Poetics' (2013) 19 *The Translator* 81.

While legal history attests to countless ways in which states have clamped down on their citizens' speech, the criminalisation of manner of expression merits critical scrutiny. A legal regime that subjects vituperation to criminalisation cannot escape the legitimacy problem by permitting a non-vituperative 'equivalent' to the discourse that is being banned. If a banned expression cannot be vituperative (and thus illegal); it also cannot be equivalent. Every speech act is constituted by its manner of expression (whether vituperate, rude, or pleasant) and the viewpoint expressed (whether racist or not).³⁷ There is no way to retain propositional content following conversion to a new medium; the very vituperation that is deemed criminalisable *only* in the context of hate speech is an indelible part of the speech targeted by hate speech bans. It is not possible to criminalise speech while permitting its expression in non-vituperative form. Hate speech bans necessarily criminalise the constitutive dimensions of the speech act itself.

It is possible to acknowledge (with Waldron) that racist speech acts may be a source of psychic or even civic harm and to agree that societies should work towards the eradication and stigmatisation of such speech, while rejecting the claim that a non-vituperative reframing of hateful speech can ever substitute for a vituperative speech act. It is further worth noting that the 'fighting words' (*HHS*, 183) exception that Waldron invokes to legitimate the banning of certain forms of speech expressly does not apply to public discourse. The question then becomes whether we are willing to violate democratic legitimacy in order to mitigate the harms of racist speech. But, as Dworkin argues, this dichotomous formulation is misleading because the privileging of the latter over the former negates the foundations of the democratic social order: it undermines the dignity owed to every citizen, and which is best embodied in the principle of fairness.³⁸ No amount of tolerance or civility can compensate for the suppression of public discourse.

I would add to Heinze's insistence that 'the citizen's prerogative of non-viewpoint-punitive expression in the public sphere is both conceptually prior to and constitutive of any legitimately democratic constitution'³⁹ that this model for free speech presupposes a democratic state actively working towards social, racial, and economic equality, and proactively challenging racism's harms. While this argument may appear circular, inasmuch as it relies on a multiplicity of conceptual priorities, accepting it does not detract from the possibility of empirically address political legitimacy retrospectively. We can claim conceptual priority for the citizen's speech prerogative, while recognising that conceptual priority rarely maps neatly onto historical sequence. The advantage of adopting a model that recognises the conceptual priority of the speech right for democratic legitimacy, while also recognising that free speech has no democratic value in a state that is not committed to social, racial, and gender equality is that it enables us to avoid the balancing model that dominates the adjudication of free speech in liberal democracies, and which more often than not serves as a justification for constraining various rights.

A speech act's manner of expression cannot be separated from its content, either for the purpose of criminal adjudication or the philosophy of language. This is why it is both accurate and necessary to insist that hate speech bans criminalise thought. Specifically,

³⁷See Rae Langton, 'Speech Acts and Unspeakable Acts' (1993) 22 *Philosophy & Public Affairs* 293, and Judith Butler, *Excitable Speech: A Politics of the Performative* (Routledge 1997).

³⁸Dworkin, 'Reply to Jeremy Waldron' ('regulating hate speech is not, after all, a matter of balancing' 342). For Dworkin's concept of dignity, see Dworkin, *Justice for Hedgehogs* (Harvard University Press 2011) 191–218.

³⁹Personal correspondence on file with the author (13 October 2018).

they criminalise thought in its capacity as *logos*, constituted by speech as well as reason, and specifically public reason, the ‘interactive process of communication’, rather than private contemplation.⁴⁰ Hate crime legislation by contrast treats racism as *mens rea*, evidence of criminal intent. While there are legitimate exceptions to the free speech principle, including the law and order concerns highlighted by Waldron,⁴¹ it is categorically not the case that hate speech bans have ‘no greater effect on political legitimacy than ... [the] acknowledged exceptions to the free-speech principle’ (183). The chilling effect imposed by hate speech bans that criminalise the constitutive linkage of a hateful viewpoint to a vituperative manner of speech cannot be compared with other exceptions to the free-speech principle, such as security concerns or the danger of inciting violence, wherein the constitutive nature of the link between viewpoint and manner of expression is absent. A security exception to the free speech principle is not made on the basis of the viewpoint expressed; by contrast, a hate speech ban is necessarily viewpoint-restrictive.

Viewpoint-based restrictions violate the prohibition on viewpoint-selective censorship that lies at the foundations of First Amendment jurisprudence since Oliver Wendell Holmes’ dissent in *Abrams v. United States* (1919), as first formally adopted by the court in *Brandenburg v. Ohio* (1969).⁴² In the present context, such restrictions also distinguish the hate in hate speech from the hate in hate crime. The aggravated penalties that accrue to hate crimes (in contradiction to criminal acts not associated with hate) do not require their reclassification as crimes. By contrast, criminalising hate speech reclassifies its status as a speech act. Through a combination of viewpoint and manner of expression, speech that is liable to criminal sanction due to its dissemination of hate is constitutively distinct from speech that is not criminally liable. Such is the difference between the hate in hate speech and the hate in hate crime.

To the extent that speech bans impose viewpoint-based restrictions, yet the views that they seek to censor are not criminalised outside the speech act, the hate speech ban proponent finds herself enmeshed in irresolvable contradictions. The point of view expressed makes hate speech subject to criminalisation, yet Waldron notes that effective hate speech bans ‘bend over backwards to ensure that there is a lawful way of expressing something like the propositional content of views that become objectionable when expressed as vituperation’ (190). On this account, even though hate speech bans are necessarily content-restrictive, effective bans contain the proviso that the same criminalised content, expressed without vituperation, is not to be treated as criminal. This argument reflects a dualistic approach to language that is out of touch with the most current thinking on the subject, as well as with the most ancient traditions of reflection concerning the verbal sign’s navigation of the relationship between form and content.⁴³

⁴⁰See further, Heinze (n 13) 105–06.

⁴¹See the example of the public order charge upheld against a distributor of antisemitic material in *The Harm in Hate Speech*, 204–7.

⁴²See *Abrams v. United States* 250 U.S. 616 (1919) and *Brandenburg v. Ohio* 395 U.S. 444 (1969), respectively. For the history of First Amendment jurisprudence, with emphasis on the Supreme Court’s relatively recent turn to free speech, see Laura Weinrib, *The Taming of Free Speech: America’s Civil Liberties Compromise* (Harvard University Press 2016).

⁴³These views have long been cornerstones of post-Saussurean linguistics and literary theory, made most famous perhaps in Jacques Derrida, *De la grammatologie* (Les Éditions de Minuit 1967). Similar insights concerning the intrinsic relation between form and content in constituting the verbal sign underwrite much of the philosophy of language and rhetorical theory in Islamic thought. See Alexander Key, *Language Between God and the Poets: Ma’na in the Eleventh Century* (University of California Press 2018).

The above points can be further illustrated by returning to the example of the speech ban embedded within the IHRA definition: when the definition's supporters use it to oppose, ban, or otherwise censor criticism of Israel, their objection generally revolves around a certain manner of (anti-Zionist) expression, which takes the form of vigorous, disproportionate, and sometimes vitriolic critique. It is not the mere criticism of Israel, such proponents say, that they wish to have classified as antisemitic, but criticism of Israel that assumes a vicious character, that involves a double standard, and which in their view is linked to contemporary antisemitism. The recent imposition of the IHRA definition in this context within UK universities and public institutions has given rise to a number of unresolved questions: who determines what is and is not antisemitic? Are such matters to be adjudicated by the victim? If there is no immediate target, who is victimised by such speech? How do we distinguish the taking of offense from the racist infliction of harm?

While such matters are difficult and arguably impossible to resolve through legal adjudication, one thing is certain: forbidding a certain manner of expression is not consistent with permitting the propositional content of its critique. We must choose. If we forbid the manner, we also forbid the expression of its propositional content. The causing of offense to passive or active supporters of Israel is the *point* of much anti-Israel discourse.⁴⁴ Events such as Israeli Apartheid Week are seen to be abrasive, offensive, and otherwise intolerable by those who wish to classify it as antisemitic and to ban it with the assistance of the IHRA definition.⁴⁵ When this discourse is censored on the grounds of its perceived antisemitic content, or due to a perceived contribution to antisemitism, it is a viewpoint (and not merely a manner of expression) that is suppressed. This viewpoint relates not only to the critique of Israel, but specifically to the appropriateness of causing offense to Israel supporters, and to anyone who seeks to minimise discussion of human rights violations in the Occupied Territories. Because the causing of offense is intrinsic to the meaning of such protest, the anti-Israel discourse that marks Israeli Apartheid Week is a mode of dissent that cannot be 'translated' into a less vituperative manner of expression without relinquishing its content.

Permitting the propositional content of a view to be expressed does not ameliorate the harms entailed in censoring the original speech act, which cannot be surgically separated from its vituperative manner. Expressive content that cannot legally be expressed in its original form undergoes an irrevocable change in character when the manner of expression is obliterated. We can therefore conclude that it is not merely the 'hate' in hate speech that is criminalised by hate speech bans, but also the manner of its expression. Yet, just as the hate in hate speech is not necessarily but only potentially criminal, so is vituperation only criminalised when it becomes a medium for hate speech. While ban proponents may claim that the fighting words exception to free speech applies here, such that it is not precisely vituperation that is criminalised but rather its security implications, I have argued that the situation is more complex. By contrast, hate crimes are not

⁴⁴Amid a dense literature on this topic, Brian Klug's reflections on causing offense within the Judaic tradition is particularly apt; see Klug, *Offence: The Jewish Case* (Seagull Books 2009).

⁴⁵For one example (of many) of an effort to use a the IHRA as a speech ban to engage in censorship, see the Simon Wiesenthal Center's press release 'Other Universities Should Follow British University's Cancellation of "Israel Apartheid Week"' (21 February 2017) <www.wiesenthal.com/site/apps/nlnet/content.aspx?c=IsKWLBpJLnF&b=8776547&ct=14985817>. The debate around the definition, including efforts to incorporate it into law, is summarised in Gould (n 24).

criminalised by virtue of the role of hate as a motivating factor. Rather, evidence of hate assists in the determining the penalty for a hate crime. These differential taxonomies bright into focus a conceptual challenge: if neither hate nor vituperation can appropriately or legitimately be subjected to criminalisation when considered separately, by what right are they suddenly subject to criminalisation when they become intertwined in public discourse?

To summarise the argument thus far: hate *plus* vituperation criminalises speech for proponents of hate speech bans. This criminalisation is illegitimate on my view because the culpability of the speaker of hate speech is distinct from that of the perpetrator of a hate crime, in which contest the *actus reus* generates criminality prior to the determination of penalties based on the discovery of hateful intent. Hate crime legislation does not infringe on the citizen's prerogative of free expression because the criminalised status of a hate crime is conceptually prior to its the secondary classification. By contrast, hate speech bans advocate for subjecting speech to criminalisation with respect to its manner as well as to its message. The contrast between hate speech and hate crime could not be starker.

The above-noted distinctions are implicit in Waldron's analysis, and lie at the foundation of the upstream/downstream distinction that he borrows from Dworkin. Yet, they tend to get buried in his refutation of Dworkin's legitimacy thesis. To the charge that racists will be excluded from expressing their views when hate speech is outlawed, Waldron avers: 'to the extent that the individuals' preferred means of expression is harmful, and ... that other means of expression are available for communicating their opposition' to hate speech bans, 'the loss of downstream legitimacy incurred as a result of the banning of speech of these particular kinds is minimal' (183). Waldron perceives no substantive risk of hate speech bans infringing on citizens' speech rights because citizens in his view always have the opportunity to voice their opposition to these bans, and to express their racism in a non-vituperative manner, without having their words subject to criminalisation.

This response to the legitimacy thesis is problematic on two grounds. First, on the matter of harm. If the purpose of banning hate speech is to reduce its harmful effects, the distinction between vituperative and non-vituperative expression as the key element in determining the appropriateness of criminalisation seems peculiar. It fails to recognise (and might prevent others from recognising) the most profound and lasting effects of racial inequality and hate-based discrimination. Against generations of structural and materialist critiques of racism, the vituperative/non-vituperative distinction problematically anchors the harms of racism in its manner of expression. The distinction thereby perpetuates a tendency within certain strands of critical race theory to advocate the censorship of 'words that wound' based on the manner of their expression.⁴⁶

Consider the term 'illegal', increasingly being applied to immigrants to the US and the UK. This epithet is non-vituperative, and is regularly used by mainstream politicians as shorthand for 'illegal migrants'. Its usage is consistent with polite discourse. Yet reducing a person to the category of 'illegal', and justifying the denial of rights to them on that basis,

⁴⁶See most famously Mari J. Matsuda, Charles R. Lawrence III, and Richard Delgado, *Words that Wound: Critical Race Theory, Assaultive Speech, and the First Amendment* (Westview Press 1993). Waldron relies primarily on Charles R. Lawrence III, 'If He Hollers Let Him Go: Regulating Racist Speech on Campus' (1990) *Duke Law Journal* 431, reprinted as chapter 3 of *Words that Wound*.

is as if not more dehumanising than the most brutal racial slur. A speech regime that makes vituperation the determining factor in assessing the harms of hateful speech is one that is incapable of recognising how the most lethal forms of racism infiltrate not just hate speech, but everyday language. This is one reason why speech ban proponents have been charged with misidentifying the harms of racist discrimination, of muting class and socio-economic analysis, and thereby of limiting the political value of critical race theory as a tool of critique.⁴⁷ Ban proponents have responded to these critiques by noting that they see bans as merely one among many tools to remedy the harms of tools, and that the symbolic value entailed in bans that target racist speech is the determining consideration.⁴⁸

The second problem with Waldron's response to the legitimacy thesis is its misidentification of the role of hate in hate speech as compared to hate crime, and its projection of the latter onto the former. Had Waldron used this argument to address the challenge to democratic legitimacy posed by hate crime legislation, his argument would have been persuasive. Given that a hate crime is a criminal act independent of the viewpoint with which it is associated, any potential challenge to legitimacy is alleviated by giving citizens the option of protesting such legislation exclusively through non-criminal means. The state cannot be expected to permit a criminal act in order to secure its legitimacy. However, Waldron's effort to refute the legitimacy thesis is specifically addressed to the criminalisation of hate speech. Inasmuch as hate speech bans focus on the hate in speech acts rather than the hate in hate crimes, the argument that works to defend hate crime legislation does not work for hate speech bans.

Waldron's refutation of Dworkin neglects the unique aspects of the criminalising process entailed in hate speech bans, which assigns to both vituperation and hate a constitutive role. Hate alone is not enough to constitute a criminal offense, nor is vituperation. Equally, however, the decision to make vituperation and hate constitutive in determining what is and is not criminalisable makes both liable to censorship in a way that hate crime legislation does not. Put simply, whereas hate crime legislation refines and extends existing criminal law, hate speech bans introduce a different modality of criminality, premised on a combination of hate and vituperation, that remains undertheorized, in part due to its innovative character.

To conclude, Waldron's refutation of Dworkin effectively establishes that neither anti-discrimination laws nor hate crime legislation undermine democratic legitimacy. Both hate crime and anti-discrimination legislation target acts that are intrinsically criminal, entirely apart from their vituperative content (or lack thereof). Discrimination and hate crimes imperil the danger and well-being of citizens, making intervention by the state not merely advisable, but positively mandated. Waldron's arguments in favour of hate speech bans can be used to bolster the case for more aggressive hate crime and anti-discrimination legislation, although this may not have been their intended use.

Dworkin's legitimacy argument against hate speech bans still remains unanswered, however, even in the aftermath of the many attempted refutations by Waldron and others. Ban proponents frequently frame the debate as one between a vision of a state

⁴⁷Henry Louis Gates, 'Critical Race Theory and Freedom of Speech' in Louis Menand (ed), *The Future of Academic Freedom* (University of Chicago Press 1996) 119–59, and Eric Heinze, 'Truth and Myth in Critical Race Theory and LatCrit: Human Rights and the Ethnocentrism of Anti-ethnocentrism' (2008) 20 *National Black Law Journal* 107–62.

⁴⁸For a critique of the symbolic defense of speech bans that further extends the illegitimacy thesis, see Heinze (n 13) 162–65.

that engages in viewpoint-selective penalties and one that neglects racism's harms.⁴⁹ This is not an adequate account of the choice we face. We can draft legislation to punish hate crime, combat discrimination, educate citizens concerning the wrongs of racism, and give marginalised groups whose dignity is injured by racist speech fora for developing persuasive and powerful critiques of racism's wrongs. All of this can be done without compromising the legitimacy of the democratic state or endorsing viewpoint-selective censorship that criminalises thought as well as speech, and which endangers not only the racists among us, but also, and more importantly, the very groups we seek to protect from harm.

As this article has made clear, my position on free speech differs from that of Waldron, who has authored the most persuasive argument in favour of hate speech bans to date. With Dworkin, Weinstein, Heinze, and Post, I hold that speech bans as such unacceptably undermine democratic legitimacy. Yet, I agree with all of these scholars that the debate around the harms of hateful speech is one worth having, not least because of the historic injustice perpetrated by racism and its ongoing harms, including the imbalances of power it has generated. Further, in addition to advocating for positions I have critiqued, Waldron has made an important intervention in this debate by premising his advocacy of speech bans on a range of limiting considerations.

'Banning hate speech should not be understood as a way of influencing a great national debate about racial or sexual equality or religious tolerance,' Waldron writes, 'nor should it be seen as a way of contributing to the ending of that debate'.⁵⁰ Ironically, the reason Waldron identifies as a bad motive for banning racist speech is precisely that claimed by many anti-racist activists on the left. To such advocates, Waldron warns of the dangers of a political agenda consumed by identity politics. 'The key to the matter,' he states, 'is not to try to extirpate offense, but to drive a wedge between offense and harm' (*HHS*, 129). These remarks, along with many other of Waldron's important insights, reveals a legal philosopher ahead of his time, while in other respects consonant with it. The path trailblazed by Waldron is one that future generations are likely to follow, if present signs are any indication.

As this analysis has suggested, my concern is that the weak points of Waldron's argument, which resonate with the left's increasing deafness to free speech considerations, will be taken forward, rather than the strengths of his argument, which caution against abuse of hate speech bans. Within this context, Waldron's criticisms of those who deploy the wrong arguments in favour of hate speech bans and who do so for the wrong reasons are as important as are his attempts to rebut the prohibitionists. Along with Dworkin and other opponents of speech bans, Waldron understands that the battle against racism cannot be won through the coercive imposition of a speech regime. The advocacy of coercive censorship that has gained traction in recent years reveals a pervasive weakness in leftist politics. Antiracist politics have attained cultural hegemony, but they still lack access to the means of production. While antiracism is not actively opposed by the liberal democratic state, all too often the state's adherence to antiracist principles has been shown to be nominal at best and hypocritical at worst.

⁴⁹Eric Heinze, 'Hate Speech and the Normative Foundations of Regulation' (2013) 9 *International Journal of Law in Context* 605.

⁵⁰Waldron (n 4) 337.

The antiracist agenda requires infrastructural support, not censorship, in order to achieve its aims. There is no need for it to engage in coercive silencing, when it can win based on accuracy, truth, and integrity.

ORCID

Rebecca Ruth Gould  <http://orcid.org/0000-0002-2198-5406>

Acknowledgements

I would like to express my deep debt to Eric Heinze for insightful comments on an earlier draft, to Riz Mokal for helpful conversations and to Nael Alqtati and Malaka Mohammed Shwaikh for assistance in obtaining figure 1.

Funding

This research has been supported by the European Union's Horizon 2020 Research and Innovation Programme under ERC-2017-STG [grant agreement no. 759346].