

# Dignity, Harm, and Hate Speech

Robert Mark Simpson

*Law and Philosophy* 32/6 (2013): 701-28

*Abstract:* This paper examines two recent contributions to the hate speech literature – by Steven Heyman and Jeremy Waldron – which seek a justification for the legal restriction of hate speech in an account of the way that hate speech infringes against people’s dignity. These analyses look beyond the first-order hurts and disadvantages suffered by the immediate targets of hate speech, and consider the prospect of hate speech sustaining complex social structures whose wide-scale operations lower the social status of members of targeted groups. In Heyman’s and Waldron’s accounts we find plausible insights into the nature of identity-based social hierarchies, and the harms that redound to subordinated people under the operations of such hierarchies. I argue, however, that both analyses are unsuccessful as justifications for the restriction of hate speech, because they do not ultimately provide reason to think that hate speech is responsible for creating or sustaining identity-based social hierarchies.

## 1. Introduction

‘Hate speech’ is a term of art in legal and political theory that is used to refer to verbal conduct – and other symbolic, communicative action – which wilfully “expresses intense antipathy towards some group or towards an individual on the basis of membership in some group”,<sup>1</sup> where the groups in question are usually those distinguished by ethnicity, religion, or sexual orientation.<sup>2</sup> Hate speech

---

<sup>1</sup> J. Angelo Corlett and Robert Francescotti, ‘Foundations of a Theory of Hate Speech’, *The Wayne Law Review* 48 (2002): 1071-1100, p. 1083.

<sup>2</sup> Note that some others define hate speech in terms of its typical uses, e.g. vilification, harassment, defamation (see Susan J. Brison, ‘The Autonomy Defense of Free Speech’, *Ethics* 108 (1998): 312-39; Onder Bakircioglu, ‘Freedom of Expression and Hate Speech’, *Tulsa Journal of Comparative & International Law* 16 (2008): 1-49), or by reference to its typical affective consequences, e.g. feelings of distress or anxiety (see

thus includes things like identity-prejudicial abuse and harassment, certain uses of slurs and epithets, some extremist political and religious speech (e.g. statements to the effect that all Muslims are terrorists, or that gay people are second-class human beings), and certain displays of 'hate symbols' (e.g. swastikas or burning crosses). We classify such activities as hate speech if, and insofar as, they convey the idea that belonging to a particular social group warrants someone's being held in or treated with contempt.

The legal status of hate speech is an absorbing issue because it lies across a point of tension in liberal democratic thought. On one hand, liberals believe that speech must be specially protected against government interference, particularly where the political character of speech would figure as part of the rationale for the government's interfering. But liberals also typically want to use the disciplinary function of the law to combat and reform identity-based social hierarchies. The twin liberal commitments to free speech and social equality thus seem to come into conflict where hate speech is at issue.

Much of the literature on this topic comes from an American legal perspective and is, as such, framed by questions about the constitutionality of restrictions on hate speech, e.g. in the form of 'campus speech codes'.<sup>3</sup> Some authors side with the courts in defending the rigid priority of First Amendment protections for free speech,<sup>4</sup> while critics argue that the First Amendment's demands must be balanced against the socially egalitarian aspirations of the Fourteenth Amendment.<sup>5</sup> Of late, however, the discourse on hate speech has evinced an increasingly cosmopolitan perspective. In Canada, the UK, Australia, and many European states,

---

Laura Beth Nielsen, *License to Harass: Law, Hierarchy, and Offensive Public Speech* (Princeton: Princeton University Press, 2004)). On the definition I favour, by contrast, speech that conveys identity-based contempt will count as hate speech, irrespective of how it is used and how it affects its targets.

<sup>3</sup> 'Campus speech codes' are regulations used by educational institutions (usually post-secondary institutions) to deter hate speech within their confines, typically using the threat of academic suspension or expulsion. For discussion of campus speech codes in US constitutional law, see for instance Nadine Strossen, "Regulating Hate Speech on Campus: A Modest Proposal?", *Duke Law Journal* 1990 (1990): 484-573; James Weinstein, "A Constitutional Roadmap to the Regulation of Campus Hate Speech", *Wayne Law Review* 38 (1991): 163-248; Thomas Peard, "Regulating Racist Speech on Campus" in Christine T. Sistare (ed.), *Civility and Its Discontents: Essays on Civic Virtue, Toleration, and Cultural Fragmentation* (Kansas: University of Kansas Press, 2004).

<sup>4</sup> See Strossen, "Regulating Hate Speech on Campus" (1990); David Richards, "Free Speech as Toleration" in Waluchow (ed.), *Free Expression: Essays in Law and Philosophy* (Oxford: Clarendon Press, 1994); James Weinstein, "An American's View of the Canadian Hate Speech Decisions" in Waluchow (ed.), *Free Expression*, (1994); L. W. Sumner, *The Hateful and the Obscene: Studies in the Limits of Free Expression* (Toronto: University of Toronto Press, 2004).

<sup>5</sup> See Richard Delgado, "Words That Wound: A Tort Action for Racial Insults, Epithets, and Name Calling", in *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment* (Boulder Colorado: Westview Press, 1993); Kathleen Mahoney, "Hate Speech: Affirmation or Contradiction of Freedom of Expression?", *University of Illinois Law Review* 3 (1996): 789-808; Delgado and Jean Stefancic, *Understanding Words That Wound* (Boulder Colorado: Westview, 2004); W. Bradley Wendel, "A Moderate Defense of Hate Speech Regulations on University Campuses", *Harvard Journal on Legislation* 41 (2004): 407-20.

modest legal restrictions on hate speech are well-established, and not as vigorously contested as in the US. It is natural to inquire, then, about how free speech and social egalitarian ideals are being negotiated in these legal regimes, and whether there is a coherent and stable justification for these regimes restricting hate speech while professing an allegiance to free speech. With these concerns in mind, a number of contemporary authors seek to identify and assess candidate justifications that might be given in favour of legal restrictions on hate speech. The best of this work goes beyond familiar platitudes about social equality and free speech, and tries to give an account of the deeper moral foundations of these mid-level political values – e.g. the goal of democratic authenticity in free speech, or the value effective individual autonomy in social equality – and an account of the ways in which these values may be promoted or imperilled in the legal restriction or toleration of hate speech.

In this paper I examine two recent and similarly-oriented contributions to the literature, which undertake this type of exploratory analysis of hate speech law. The authors, Steven J. Heyman, and Jeremy Waldron, aim to describe part of the foundations of social equality as a political value, and in this way, to articulate candidate justifications for the legal restriction of hate speech. The pieces are thematically unified by the emphasis they place on the idea of *dignity*. Both authors regard properly formulated restrictions on hate speech as tools for securing a dig-nitarian social order.<sup>6</sup>

I begin by explaining why a harm-prevention framework – in which our principal aims are to characterise the harms that may be caused by hate speech, and ascertain whether hate speech does in fact cause those harms – should be adopted for assessing the legitimacy of legal restrictions on hate speech. I then examine Heyman’s and Waldron’s accounts in light of this framework. The central question, a response to which I look for in their work, is whether we can reasonably regard hate speech as something that effects a harmful infringement of its targets’ dignity. In Heyman and Waldron we find *prima facie* credible insights into the nature of identity-prejudicial social hierarchies, and the character of the harms that redound to subordinated people under the operations of such hierarchies. I will argue, however, that both analyses are unpersuasive as putative justifications for the restriction of hate speech, because they do not in the end give us good reasons

---

<sup>6</sup> Note, however, that neither author purports to establish a decisive case in favour of legal restrictions on hate speech. Heyman argues that hate speech should not enjoy special protection under the US Constitution, and thus he challenges a particular barrier to the legal restriction of hate speech in US courts (see “Hate Speech, Public Discourse, and the First Amendment”, in Ivan Hare and James Weinstein (eds), *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009), p. 159). Waldron, in his most recent work, says he is trying to offer a characterisation of anti-hate speech law, as something that is already a well-established part of Western liberal legal systems outside the US, in the hope that this characterisation will make it evident to American critics of anti-hate speech laws on what grounds those laws are, or might reasonably be, motivated. See Waldron, *The Harm in Hate Speech* (Cambridge Massachusetts: Harvard University Press, 2012), pp. 11-17.

to think that hate speech contributes to, and/or bears responsibility for, the establishment and perpetuation of identity-prejudicial social hierarchies, and the harms and disadvantages that individuals experience as a consequence of those hierarchies.

## 2. Harm-prevention and the restriction of hate speech

What legitimate reasons might a government have to legally restrict hate speech? Depending on one's views about the moral foundations of the criminal law, an answer to this question might advert to (i) the aim of preventing or mitigating *harms* to persons, or (ii) the aim of preventing and deterring *wrongful* acts, irrespective of their harmfulness, or (iii) the aim of affording just *deserts* to those who perform harmful or wrongful acts, or (iv) the benefits or intrinsic goods that reside in the *expressive function* of hate speech's criminalisation. If we broaden the question to include the prospect of making hate speech liable to penalty under private law, answers might advert to (v) a *compensatory* aim (damages may be levied to compensate for the distress or disadvantages caused by hate speech), or (vi) a *distributive* aim (damages may be levied to more fairly distribute the social burdens created by hate speech). And if we step back again, turning from a jurisprudential perspective to a more foundational, political-axiological view, further candidate justifications emerge; one might invoke (vii) an *emancipatory* rationale (restricting hate speech furthers the liberation of oppressed people), or (viii) a *civic* rationale (restricting hate speech encourages healthy relationships between citizens). Given this array of possibilities, and given the perennial nature of debates about the proper purposes of legal institutions, it may seem that there is little hope of assessing *the* case for legal restrictions on hate speech. Any account of that case will seemingly, of necessity, be confined to a controversial theoretical standpoint vis-à-vis the moral foundations of law.

So far as this is our situation, I believe the way forward is to assess the case for restricting hate speech under a harm-prevention paradigm. This is because no other candidate rationale for restricting hate speech becomes compelling until we have established that hate speech inflicts harm on its targets. We can take it as a given that hate speech is morally benighted and (often) profoundly offensive. But if that is *all* we can say about the adverse character and effects of hate speech, the other putative rationales for its restrictions seem relatively tenuous. We may allow (i) that it is *wrong* to behave rudely, (ii) that reactionary ideas *deserve* a reprimand, (iii) that we have good reasons to formally *express* our opposition to gratuitously offensive speech, (iv) that we have good reasons to *compensate* hurt feelings, (v) that we ought to fairly *distribute* the burdens of incivility, or (vi) that persistent, gratuitous offensiveness can be a kind of *oppression* that people would rightly want to avoid. But at most, these things seem like defeasible *pro tanto* reasons for

restricting hate speech – reasons liable to be overridden by considerations that countervail against legal restrictions on *any* conduct (e.g. costliness, risk of inefficacy, risk of sinister misuse), quite apart from the countervailing free speech concerns that specially apply in this arena.<sup>7</sup> On the other hand, if we can establish that hate speech does harm its targets, these other rationales become rather more compelling. Our reasons for legislating to remedy wrongs, mete out deserts, and distribute burdens, for instance, seem weightier when the adverse states of affairs, from which the reasons derive, involve the infliction of harm rather than offence or affront.<sup>8</sup> If the harmfulness of hate speech can be established, it remains a further question how to best prosecute the case in favour of legal restriction; for example, whether to emphasise the aim of affording just deserts to individuals who inflict harm via hate speech, over the aim of harm-prevention *per se*. My claim is that no convincing rationale for legally restricting hate speech comes into view until we establish that hate speech is harmful to its targets.

Within a harm-prevention framework, we will need to distinguish between ‘direct’ and ‘indirect’ harms that may be attributed to hate speech. No one would pretend that harm can *never* be inflicted by speech that expresses intense antipathy towards people on the basis of their race, or religion, or sexual preference. Here is an unambiguous (if abnormal) example: person A shouts homophobic abuse into person B’s ear, thereby perforating B’s ear drum. Here is a less peculiar, but similarly unambiguous example: A uses homophobic abuse to threaten and harass B over a period of months in their workplace, thereby causing B intense psychological distress. In such cases we see that acts of hate speech obviously *can* inflict harm. But this doesn’t get us very far with respect to the question of whether we should have general legal restrictions on hate speech. Where hate speech is defined by its locutionary content, to say that hate speech can inflict harm is just to say that hateful locutions may be put to harmful illocutionary uses, or that they may have harmful perlocutionary consequences.<sup>9</sup> We might have

---

<sup>7</sup> Here I do not address the questions of why and how speech should benefit from special legislative protections. Among the vast literature, some of the more original contemporary treatments of these questions include Eric Barendt, *Freedom of Speech* (2<sup>nd</sup> edn.; Oxford: Oxford University Press, 2005); Sumner, *The Hateful and the Obscene*; and James Weinstein, *Hate Speech, Pornography, and the Radical Attack on Free Speech Doctrine* (Boulder Colorado: Westview Press, 1999).

<sup>8</sup> It is a complex question how we distinguish hate speech’s harmful effects of from its merely adverse effects. I will assume here that such a distinction is tenable in principle. If hate speech perpetuates identity-based social hierarchies, as some authors claim, then the severely adverse effects associated with this can be distinguished from the offence that someone feels at the use of a slur or epithet. My interest is in whether hate speech can be ascribed responsibility for social hierarchies and their harmful effects, rather than the precise basis of the distinction between these harms and hate speech’s other, merely adverse, effects.

<sup>9</sup> The operative distinction here is from J. L. Austin, *How to Do Things with Words* (2<sup>nd</sup> edn.; Cambridge Massachusetts: Harvard University Press, 1975), between three kinds of acts that may be (simultaneously) performed in an utterance: (i) a *locutionary* act, consisting in an utterance of meaningful words, (ii) an *illocutionary* act, i.e. the act constituted by an utterance, or performed *in* uttering an utterance, and (iii) a *perlocutionary* act: the bringing about of certain effects *via* an utterance.

good grounds for restricting threats, and concomitantly, good grounds for restricting hate speech that is used to threaten, where the specific case falls under the general principle. However, the question is whether we have good reasons to restrict an instance of hate speech by virtue of its *being* hate speech, without reference to the illocutionary uses to which it is put, or to its imminent perlocutionary consequences.

The notion that hate speech inflicts harm generally, in a way that could speak in favour of this approach, may seem like an extravagant and implausible conjecture; a claim to the effect that any expressive act which qualifies as hate speech, due to its content, will necessarily harm some individual. But no authors who argue for general legal restrictions on hate speech (as far as I'm aware) make such a claim. Rather, those who hold this view of hate speech tend to think of individual acts of hate speech operating in a way that is analogous to pollution. Individual acts of pollution *can* inflict discrete harm on specifiable victims. Many acts of pollution don't inflict harm in that way. However, even when there are no specifiable victims, all acts of pollution have a degrading impact on environmental systems whose degradation beyond a certain point does inflict harms on individuals. In this way, all the polluting acts that don't directly harm anyone can nevertheless be adjudged indirectly harmful. Analogously, acts of hate speech do not always directly harm specifiable individuals, but they all contribute, so one may argue, to the creation or sustenance of a social climate in which harms and disadvantages redound to members of vulnerable social groups.<sup>10</sup> Our questions, then, will be about the extent of the contribution that hate speech makes to harmful social hierarchies, and about how its contribution compares to other contributing factors. In arguing for general legal restrictions on hate speech under the harm-prevention framework, the aim should be to show that among the many factors which operate jointly to effect identity-based social hierarchies, hate speech is not just a peripheral causal component, or something that eventuates as a downstream consequence of identity-based social hierarchies, but that it is, rather, one of the contributing causal factors through which such hierarchies are created and sustained.

---

<sup>10</sup> For an elaboration of this kind of distinction, see Joel Feinberg, *Social Philosophy* (Englewood Cliffs: Prentice-Hall, 1973), pp. 25ff; Joshua Cohen, "Freedom of Expression", *Philosophy & Public Affairs* 22 (1993): 207-63, p. 231).

### 3. Dignity, anti-hate-speech law, and the ambiguity of ‘infringement’

Several authors in the hate speech literature invoke the idea of dignity in their work.<sup>11</sup> What is often missing is an explanation of why dignity is a specially pertinent moral concept in an inquiry into the grounds for restricting hate speech. If we are trying to characterise the harms of hate speech, why not focus on the targets’ welfare, and ask whether hate speech corrodes it? Why not consider the targets’ rights, and ask whether hate speech violates them? Perhaps there are pragmatic reasons why some theorists avoid these approaches. If the adverse effects of hate speech, when conceived in welfarist or deontological terms, do not seem severe enough to underwrite an argument for restricting hate speech, committed opponents of hate speech may be prompted to retreat to some alternative moral terminology. But I do not think it is opportunism after all that orients authors in this arena towards the concept of dignity. Even if hate speech does corrode welfare and invade rights, it may still be that the best way to summarily characterise these things is to say that hate speech infringes its targets’ dignity. This is because dignity is a moral concept that is specially associated with considerations of rank, status, and hierarchy. A person’s dignity is not just a matter of her being treated well or poorly; it is a matter of how she is treated in light of the attitudes that prevail in her community, including her own attitudes, about her worth, standing, or esteem, relative to others, and relative to the community as a whole. And as I suggested above, the best way to make sense of the claim that hate speech in general inflicts harm on people, is to think of hate speech as something that contributes to (identity-based) social hierarchies.

Now, one may point out that contemporary moral discourse is typically, and appropriately, uneasy about notions of rank and hierarchy. For many it is an axiom of moral theorising that there is, or should be, no hierarchical rankings amongst moral subjects: that no one’s happiness or preferences count for more than any others’ (for certain utilitarians), or that everyone is equally endowed with rights in virtue of their humanity (for certain Kantians). How, then, can dignity occupy

---

<sup>11</sup> For example: “the most stringent demands of... respect for the dignity of persons... pull strongly in favour of formally uniform hate speech rules” (George R. Wright, ‘Dignity and Conflicts of Constitutional Values: The Case of Free Speech and Equal Protection’, *San Diego Law Review* 43 (2006): 527-76, p. 569); “a society that permits toleration to run roughshod over human dignity deceives itself into thinking that undeserved abusive expression is worth the cost of denying many persons what is basic and decent” (Corlett and Francescotti, ‘Foundations of a Theory of Hate Speech’ (2002), p. 1097); “a racial insult is always a dignitary affront... a direct violation of the victim’s right to be treated respectfully” (Richard Delgado and Jean Stefancic, *Must We Defend Nazis? Hate Speech, Pornography, and the New First Amendment* (New York: New York University Press, 1997), p. 8). We also find appeals to dignity in anti-hate speech legislation, for example in Germany, where “the Federal Constitutional Court has ruled that the Basic Law establishes ‘an objective ordering of values’ with some rights being more important than others” so that, “freedom of speech, press, assembly, and association are... inferior to the government’s interest in securing and protecting human dignity” (Ronald J. Krotoszynski, ‘A Comparative Perspective on the First Amendment: Free Speech, Militant Democracy, and the Primacy of Dignity as a Preferred Constitutional Value in Germany’, *Tulane Law Review* 78 (2004): 1549-1609, pp. 1562-63).

a legitimate place in an egalitarian moral mindset? The answer is that while the concept of dignity does indeed connote some sense of conferring status-based honour and respect, via adherence to customary modes of interaction, the contemporary understanding of dignity demands that honour and esteem be portioned out to all in equal measure, insofar as we all occupy an equivalent rank, as members of a shared moral community.<sup>12</sup> An infringement of person A's dignity, for us, does not mean failing to treat A in accordance with the demands of a caste system. On the contrary, it means treating A as though she *does* belong in a caste system, and thus failing to treat her as a fully-entitled person or member of society.

Because the concept of dignity orients our attention towards considerations of rank, status, and hierarchy, an appeal to this concept opens up a moral vocabulary that is especially useful in the discussion of hate speech. We cannot properly apprehend what is wrong about expressing intense antipathy towards gay people, indigenous national groups, immigrant communities, religious minority groups, and others, if we think of the hate speaker and his targets as opponents engaged in an isolated, private conflict, in which one party's hostility to the other is the only concern. Rather, so I want to suggest, we should recognise that the targets of identity-based hate speech are vying to secure a position of esteem and recognition within a wider social and cultural ecosystem, and that this aspiration is what is being opposed – both maligned, and also, some authors claim, impaired – in the hate speaker's verbal conduct.

That is the principal reason to favour a dignity-based approach, at any rate. At the same time, though, dignitarian accounts of the harm of hate speech are difficult to appraise, because of the ambiguity in the language that is available to us for describing actions or policies that stand in either a negative or positive relationship to people's dignity. Dignity may be 'infringed', 'violated', 'invaded', 'assaulted', and 'trampled-upon'; or it may be 'respected', 'recognised', 'protected', and 'upheld'. When person A fires a demeaning insult at person B in a public arena, there is obviously some sense in which A's speech 'goes against' B's dignity. At a minimum, A advertises to B and others her own refusal to afford B the respect that B is owed by virtue of her status as a fellow citizen or person. But how should we characterise the antagonistic relationship between A's insult and B's dignity, beyond this observation? The answer cannot be ascertained just by reflecting on the content and character of the speech in isolation, because it is a further question whether A's insult will diminish B's status in the eyes of B's peers, or of B herself; and if B's esteem remains intact, then it will at least be overwrought, if not misleading, to say that B's dignity has been 'assaulted' or 'trampled-upon' through A's speech. Our problem, stated more generally, is that there is an ambiguity between two senses in which a person's dignity may be acted against. One

---

<sup>12</sup> See Michael Walzer, *Spheres of Justice: A Defence of Pluralism and Equality* (Oxford: Basil Blackwell, 1983), pp. 252-53.

sense is related to the actor's attitudes: A thinks of B as a second-class person, and she expresses that view to others, or tries to influence others towards accepting that view for themselves. The other is a matter of social outcomes: A actually succeeds in diminishing B's status in their shared community. Words like 'infringe', 'trample', and 'assault' are ambiguous between these two ways in which a speaker might 'act against' another person's dignity, and terms like 'respect' and 'protect' are ambiguous between the two sorts of dignity-oriented 'infringements' that they might be aiming to mitigate or redress.

Some authors seem to regard it as uncontroversial that hate speech infringes against its targets' dignity. This judgement is entirely plausible if one thinks of infringing dignity in the first sense. However, as I suggested in §2, the case for general legal restrictions on hate speech hinges on our establishing that hate speech is responsible for inflicting social-hierarchy-based harms on its targets. In order to vindicate a dignity-based argument for restricting hate speech, under this framework, we need to see that hate speech infringes its targets' dignity in the second way: that it diminishes their status in some discernible sense, beyond merely being a manifestation of the speaker's view of them as inferior or second-class people.

#### 4. Heyman on status and recognition

Heyman's discussion of hate speech and dignity plunges into these difficult waters.<sup>13</sup> On Heyman's construal, dignity consists in a person receiving from others a particular form of recognition, one which underlies the person's status as a rights-bearer generally. "Rights are rooted in respect for personhood" he says, and this means "an individual cannot enjoy rights in relation to others unless they recognise him as a person".<sup>14</sup> An emphasis on recognition in moral and political theory is not at all eccentric. In the German Idealist tradition, Hegel, most famously, emphasises recognition – roughly, the apprehension of self-government in others, and the redirecting of one's own intentions and desires in light of this apprehension – as the foundation of ethical life and thought.<sup>15</sup> Among contemporary political theorists, Habermas defends a Hegelian ideal of recognition as a counterpoint to the prevalent ethical frameworks of modernity, with their tendency (so Haber-

---

<sup>13</sup> Here I am focussing on Heyman's account in "Hate Speech, Public Discourse, and the First Amendment" (see note 6). This expands upon an earlier account in Heyman's *Free Speech and Human Dignity* (New Haven: Yale University Press, 2008). Page references to Heyman's work, in this and the following section, all refer to the later article.

<sup>14</sup> pp. 166-67.

<sup>15</sup> See Robert R. Williams, *Hegel's Ethics of Recognition* (Berkeley: University of California Press, 1997).

mas contends) to favour introspective attention to one's own preferences and desires as the pathway into ethical understanding.<sup>16</sup> For other modern authors, an emphasis on recognition presents one way to try to reconcile an individualistic view of rights with the need for respecting diversity that arises in culturally heterogeneous societies.<sup>17</sup> I mention these points just to acknowledge that Heyman's account is situated within a rich theoretical tradition. What is distinctive about his work is that it calls for the enshrinement of a 'right to recognition' via legal obligations for individuals. It is not only governments and institutions that are obliged to recognise people's status or identity, and treat them with respect accordingly. Rather, Heyman argues, these duties of respect obtain for *all others* in relation to whom the subject's rights obtain. In order for person A's right to recognition to be respected, A must be recognised as a full person or citizen by everyone else in her community, and treated by all in accordance with that status. Heyman characterises the harm of hate speech as a violation of this putative right to recognition. By calling for the exclusion or oppression of A's ethnic group (for instance), or by saying that A's group are contemptible, second-class persons, the hate speaker signals his refusal to accord A and others the recognition they are owed. The hate speaker's non-recognition, so Heyman argues, diminishes the targets' social standing and thereby harms them.

One way to think about this view is to see it as part of a move away from institutional preoccupations in liberal egalitarian thought. Will Kymlicka reminds us that "until the 1960s, the state itself was still engaged in explicit discrimination against women, homosexuals, racial minorities, and so on", and that many expected the end of official state discrimination to usher in an egalitarian milieu for historically subordinated groups.<sup>18</sup> The disappointment of these expectations in the post-civil rights era has led many liberals to the view that a society which aspires to genuine equality for its members cannot content itself with formally equitable institutions. Governments must seek to influence civil society in a way

---

<sup>16</sup> Jürgen Habermas, *The Philosophical Discourse of Modernity* (trans. F. Lawrence; Cambridge Massachusetts: MIT Press, 1987), pp. 41ff.

<sup>17</sup> See Walzer, *Spheres of Justice* (1983), pp. 258ff. Or consider the following from Charles Taylor: "The thesis is that our identity is partly shaped by recognition or its absence... and so a person or group of people can suffer real damage, real distortion, if the people or society around them mirror back to them a confining or demeaning or contemptible picture of themselves" (Taylor, *Multiculturalism And 'The Politics of Recognition'* (Princeton: Princeton University Press, 1992), p. 25). Note, however, that the recognition Heyman says people are owed is not identical with the recognition whose deprivation inflicts harm, on the view Taylor describes. As Steven Rockefeller says, and Heyman would agree, "from the liberal democratic point of view, a person has a right to claim equal recognition first and foremost on the basis of his or her universal human identity and potential, and not primarily on the basis of an ethnic identity" (Rockefeller, "Comment" in A. Gutmann (ed.), *Multiculturalism And 'The Politics of Recognition'* (Princeton: Princeton University Press, 1992), p. 88).

<sup>18</sup> Will Kymlicka, "Civil Society and Government: A Liberal-Egalitarian Perspective" in Nancy L. Rosenblum and Robert C. Post (eds.), *Civil Society and Government* (Princeton: Princeton University Press, 2002), pp. 102-103.

that goes beyond formal institutional structures, so that the egalitarian aspirations that shape and govern our institutions do not ultimately remain confined to those locations. Heyman's legal right to recognition can be seen as a tool for mobilising such influence. His approach also shares some features with another reformist trend in liberal theorising, namely, the move away from conceiving of citizenship as a set of status-based entitlements, towards seeing citizenship as encompassing both rights *and* duties.<sup>19</sup> On this conception of citizenship, the right to freedom of speech might carry with it a responsibility not to jeopardise or undermine, in one's speech, the civic bonds that hold society together.

Nevertheless, one could accept these ideas while rejecting the legal duties posited in Heyman's account. When political theory is preoccupied with the rights of individuals against governments, it can ignore the ways in which a system which respects *those* rights may fail to secure other legitimate claims for its positionally disadvantaged members, and why further measures – encouraging responsible citizenship, and promoting *de facto* equality – are required. But Heyman's view goes further. His contention is that people's civic statuses are sensitive to, and indeed can be diminished by, the verbal mistreatment that people receive at the hands of hate speakers. Most liberals will accept that 'hate speech does not respect others' equality or dignity', but they will see the crucial consideration that is 'at stake in evaluating the content of the legal order' as the respect or lack of respect that is demonstrated by *the government*, rather than by hate speakers and others.<sup>20</sup> Heyman seemingly purports to deny the normative significance of that distinction. Rights of recognition "are founded on the duty to respect the autonomy and dignity of human beings" and this, he says is "a duty that applies not merely to the state but also to individuals".<sup>21</sup> Speech whose content is contemptuous can diminish the target's civic status, even when the speech in question does not take place within the context of a government institution, and even when it does not represent the values of, and is not made on behalf of, the community at large.

Taken to its extreme, Heyman's argument would efface the distinction between harms that individuals inflict on each other, and harms that governments and social institutions inflict on individuals. My neighbour may beat me. The state may disenfranchise me. Both experiences make me worse off; the question is whether they make me worse off in different respects, which, in principle, call for different legal remedies. When person A inflicts harm on person B by beating him, we naturally want to say that *A and A alone* is the one who harms B. B would be mistaken

---

<sup>19</sup> See Will Kymlicka and Wayne Norman, "The Return of the Citizen: A Survey of Recent Work on Citizenship Theory", *Ethics* 104 (1994): 352-81.

<sup>20</sup> Edwin C. Baker, "Autonomy and Hate Speech" in Ivan Hare and James Weinstein (eds.), *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2009), p. 143.

<sup>21</sup> p. 176.

if he saw himself as having been mistreated by the state or by society at large, as an *ipso facto* consequence of his being beaten by A. Unlike the person who is disenfranchised, or structurally discriminated against in the education system, the abuse of B does not constitute a diminution of B's civic status. The cases in which an individual *can* claim mistreatment by the state or society at large, according to this view, are just those cases in which the individual is ascribed an inferior status in the institutions through which his participation in society is unavoidably mediated, such as the legal system, the political system, and the education system.<sup>22</sup> Part of what Heyman is doing is challenging this way of thinking, at least as far as it applied to hate speech. He suggests that hate speech performed by private citizens can diminish people's status in the way we standardly associate with an abusive relationship between the state and the individual. People don't just hold their right to recognition against the state and its institutions, they hold it against society *per se*, and consequently hate speech from any individual may be a rights-violating incursion of society against the individual.<sup>23</sup>

## 5. Formal and informal dimensions of equality and status

Something that is unclear in Heyman's account is whether he believes there is a distinction to be drawn between a person's *actual* civic status (however exactly we define that), and the civic status that the person is implicitly ascribed in the treatment she receives from peers and neighbours. This unclarity is important to seeing why his argument is ultimately unconvincing. An example will be useful

---

<sup>22</sup> The contrast I am drawing between B being mistreated by his neighbour, A, and B being mistreated by the state, is somewhat over-simple. A's maleficent attitude towards B may have been conditioned by the state's discriminating against (or tolerating discrimination against) B's ethnic group. A may know, or reasonably expect, that he is unlikely to be arrested or prosecuted for attacking B, due to structural discrimination against B's ethnic group in the legal system. In both cases, the state bears some responsibility for the harm A does to B. My point is that in cases where such conditions do not obtain, there is a distinction – which Heyman calls into question – between harms enacted in the relations between persons, and harms enacted in the relations between the state and its subjects. Thanks to an anonymous referee from this journal for pressing me on this point.

<sup>23</sup> It is somewhat ambiguous whether the negative character of these incursions, on Heyman's view, ought to be elaborated in the language of harm, or whether he is just making a more modest claim about harmless rights-violations. Heyman frames his discussion as a rejoinder to the claim that hate speech should be protected as free speech; his response is (roughly) that it should not receive special protection, since it violates people's right to recognition, and since free speech protections are circumscribed (so he argues) by how the rights on which they are based interact with other rights. At other points, though (e.g. at p. 166, and again at p. 179), Heyman seems to be crafting a positive case for legal restrictions on hate speech, the argument being that hate speech has a harmful effect on its targets' dignity, and should therefore be prohibited. I have interpreted Heyman's account along the lines of the second reading, since, for the reasons I state in §2, this is the kind of argument one would need to justify restrictions on hate speech all things considered (as opposed to what one would need to defend those restrictions against free-speech-based objections). I should acknowledge, however, that the success of the more modest (rights-based) argument, that hate speech should not be *immune* from legal restriction under the auspices of a free speech principle, is independent of any judgement about the merits of the more ambitious (harm-based) argument in *support* of legal restrictions on hate speech.

here. In modern liberal democracies, mixed-ethnicity couples have the same rights and entitlements as others, both individually and with regards to the status of their relationship. Of course this was not the case two centuries ago, and there remains a small minority in modern liberal societies who would prefer the legal and cultural norms of the early-19<sup>th</sup> century to remain in effect in this area. Now, suppose we have an advocate of ‘traditional values’ in this area, A, who is so zealous in his views that he treats a particular mixed-ethnicity couple with whom he interacts – in their mutual workplace, let’s say – the way he believes (and anyone else who shares his 19<sup>th</sup> century beliefs and convictions *would* believe) they deserve to be treated. Thus, A regularly voices his indignation about this couples’ relationship, and he does what he can within his narrow sphere of influence to spoil their participation in civil society, and their ability to enjoy the benefits of that participation. In short, A treats the couple as though they are inferior on account of their relationship, and their social identity to the extent that it is bound up in that relationship. Now, however appalling A’s conduct, the imagined social hierarchy that underwrites it does not, and without a seismic shift in the social order could not, determine the social standing of this or any mixed-ethnicity couple. When A treats the couple as second-class citizens, there is a sense in which, in addition to his gross moral failings, he is simply factually mistaken. The equal civic status of mixed-ethnicity couples in modern liberal societies is enshrined both in anti-discrimination legislation, and in social practices which largely (though of course imperfectly) affirm the ideals that underpin such legislation. The judgement that the couple’s civic status is *non*-subordinate, in the contemporary context, is based as much upon the latter – there being support for the relevant norms of anti-discrimination – as it is on the fact that those norms carry the special imprimatur of being legally enforceable across certain domains.<sup>24</sup>

In rejecting the idea that formal legal apparatuses are the sole determinant of a person’s civic status, a defender of Heyman’s account may be tempted to say that there is no fact of the matter as to how one person’s civic status compares to another person’s; that there is no abstract truth about what person B’s civic status amounts to, only facts about how other people treat B. But this cannot be what Heyman thinks, because if there are no stable facts about people’s civic statuses, there is no intelligible reason for singling out ‘violations’ of people’s putative right to recognition for condemnation. The converse temptation, then, would be to say that because people’s statuses are partly informal, grounded in the beliefs, attitudes, and expectations of the wider populace, someone’s being informally treated as a second-class citizen – e.g. by being the target of hate speech – means

---

<sup>24</sup> This distinction appears to have been overlooked in Andrew Altman’s arguments about US campus speech codes (‘Liberalism and Campus Hate Speech: A Philosophical Examination’, *Ethics* 103 (1993): 302-17). For Altman, hate speech constitutes the illocutionary act of “treating someone as a moral subordinate” (p. 310), and this is why it should be restricted in educational environments. But he fails to acknowledge a difference between something being a ‘speech act of subordination’ (i.e. something that *actually causes* subordination), and a speech act that merely treats someone *as though* they are morally subordinate.

that, *ipso facto*, they have been made a second-class citizen. But this conclusion assigns too much significance to the attitudes and conduct of noisy, marginal individuals. When a political and legal regime is egalitarian only 'on paper', and in reality is pervaded by identity-based prejudices and hierarchical, discriminatory practices, then it may indeed be the case that members of vulnerable groups have an inferior civic status, their *de jure* equality notwithstanding. However, if a group's equal standing is affirmed both by the law *and* by informal social norms (as in the mixed-ethnicity couple's situation), then merely being spoken of as socially inferior by an isolated bigot does not by itself confer a subordinate status upon members of the relevant group.

The view that I am attributing to Heyman, when duly qualified, does contain a grain of truth. We will see the damaging potential of hate speech in a different light if we think of people's civic status as something constituted by the respect and recognition people receive in all their social relations, and not only their institutionally-mediated relationship with the state. However, even when we are thinking about people's status from this perspective, what we are faced with in modern liberal societies is a complex array of social forces, which includes manifestations of identity-based prejudice, but which also includes many forces that reject and counteract those identity-prejudicial elements – including a (largely) functional and (largely) well-supported legal system, which purports to identify, mitigate, and redress identity-prejudicial practices. My claim is that when we are trying to appraise this sort of 'mixed system', it seems odd that the conduct of hate speakers should be seen as a key factor, out of the whole multifarious jumble, in determining the social status of people in vulnerable groups. In saying this I don't mean to suggest that identity-based prejudice has been consigned to history's dustbin by contemporary anti-discrimination law. To take my previous example, mixed-ethnicity couples may still be treated as *de facto* second-class citizens in certain notionally liberal jurisdictions. And even in the most progressive legal regimes, there are still certain markers of identity on the basis of which people are made subordinate, in a variety of formal and informal ways. Nor do I mean to suggest that there is no relationship at all between hate speech and systemic identity-prejudice. The agential oppression carried out by the hate speaker is in some sense aligned with, and is thus liable to reinforce, the structural oppression that people in vulnerable groups experience.<sup>25</sup> My concern is with the degree of blame and responsibility we can reasonably ascribe to hate speakers. Where *de jure* and *de facto* social equality both largely obtain, and reinforce one another, the hate speaker's dissident voice by itself is insufficient to derail this and diminish people's status. And where *de jure* equality co-exists with various forms of *de facto*

---

<sup>25</sup> I borrow the distinction between agential and structural oppression from Sally Haslanger, "Oppressions: Racial and Other" in Michael P. Levine and Tamas Pataki (eds.), *Racism in Mind* (Ithaca: Cornell University Press, 2004), pp. 104-05.

inequality, we have seen no reason, apart from the mere fact of 'ideological alignment', to suppose that hate speech is responsible for this wider *de facto* inequality. We can grant, again, that the hate speaker assaults his targets' dignity in the individualistic sense that he personally fails to accord them the respect and recognition they are owed. But whether or not they are accorded these things more generally, by their wider community, is not, as far as anything Heyman has shown us, determined by the individual hate speaker's verbal conduct.

## 6. Waldron on equality and assurance

I am arguing that in order for the hate speaker to be made answerable for infringing his targets' dignity, it must be the case that his speech contributes to the targets' inferior social status, not just in the microcosmic domain of the antagonism between speaker and target, but in the wider social realm they both inhabit. Given the numerous elements that operate together to determine a person's status at this wider level, including many formal institutions and countless unstructured social practices, it seems improbable that the sentiments voiced by a relatively disorganised rabble of hate speakers could be *the*, or even *an*, important factor in this respect. Certainly we cannot reach that conclusion just by observing that hate speakers speak out in opposition to their targets' dignity, thereby signalling their own unwillingness to treat their targets as equals, and implicitly calling on others to follow suit.

Perhaps, however, in contrasting the hate-speech-target's status as it actually stands in the wider social arena (which is the key consideration from a legal perspective), and the target's status in the eyes of the hate speaker (which does not determine the target's status, beyond the speaker's own estimations), we overlook the significance of how hate speech affects its targets' perceptions of their own social status. Waldron argues that hate speech infringes the dignity of its targets via this mechanism: not by degrading their equal status as it is enshrined in law or manifested in egalitarian attitudes and practices, but by undermining the targets' *assurance* regarding the security of their dignitarian status.<sup>26</sup>

Waldron conceives of hate speech as a type of group defamation. This seems strange at face value. Social groups are not legal persons, and hence the lowering of a group's reputation is not a legally actionable harm, except perhaps indirectly,

---

<sup>26</sup> Here I focus on the analysis and defence of anti-hate speech laws that Waldron develops in 'Dignity and Defamation: The Visibility of Hate' (*Harvard Law Review* 123 (2010): 1596-657). The arguments in question have been advanced in several places. They first appeared, more embryonically, in Waldron's 'Free Speech and the Menace of Hysteria', *The New York Review of Books* (volume 55, 2008). The "Dignity and Defamation" article is a published version of the 2009 Oliver Wendell Holmes lectures, which were delivered at Harvard Law School, partly in response to the reception Waldron's earlier article received. The arguments have since reappeared, with minor alterations and additional discussion of surrounding issues, in *The Harm in Hate Speech* (2012) (see note 6). Page references to Waldron's work in this and the following section refer to the 2010 article, apart from one quotation which I clearly identify.

since groups are comprised of individuals who may suffer reputational damage.<sup>27</sup> We can imagine a defamation-based anti-hate speech law modelled on the tort of libel, with the hate speaker being liable to penalty for impugning people's reputations *by* impugning the reputation of the group to which they belong. But whatever merits this approach may have, Waldron takes an alternative route, endorsing a conception of group defamation modelled on criminal libel statutes.<sup>28</sup> Such laws have latterly fallen out of favour in modern legal systems, but their onetime purpose, for Waldron, provides an instructive model for our thinking about hate speech law. Unlike civil libel statutes, which empower individuals to protect their reputations against personal attack, criminal libel statutes were a way for states to preserve public order, and thereby foster the necessary conditions for the maintenance of people's reputations. But how can libellous statements corrode public order, and how can the criminalisation of libel protect public order? One possibility is that libellers run the risk of creating disorder through stimulating the insulted parties to revenge, and that criminalising libel forestalls this risk.<sup>29</sup> But Waldron draws our attention to a deeper relationship between group libel and public order. Criminal libel laws protect something that public order depends upon: a widely-shared understanding and acceptance among the populace of the status-respect that we all owe, and are owed, as fellow members of society. Libellous statements can function as an attack on the acceptance of our mutual standing *qua* citizens, on which public order in turn relies.<sup>30</sup> And thus, Waldron suggests, restrictions on hate speech can be seen as operating in the same way as criminal libel statutes once did, as a way of protecting these dignitarian foundations of public order.<sup>31</sup> Hate speech regulation, he says,

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

---

<sup>27</sup> Note, however, that some theorists have nevertheless argued for narrowly-drawn legal restrictions on hate speech primarily on the basis of harms to social groups (e.g. Thomas David Jones, *Human Rights: Group Defamation, Freedom of Expression, and the Law of Nations* (The Hague: Kluwer Law International, 1998)).

<sup>28</sup> A broadly similar approach to Waldron's was explored in earlier work by Hadley Arkes, 'A Communitarian Defense of Group Libel Laws', *Harvard Law Review* 101 (1988): 682-701.

<sup>29</sup> For an argument to this effect see Graham Hughes, 'Prohibiting Incitement to Racial Hatred', *The University of Toronto Law Journal* 16 (1966): 361-68.

<sup>30</sup> pp. 1604-05.

<sup>31</sup> The idea that hate speech restrictions protect public order has been advanced by others (e.g. Alexander Tsesis, 'Dignity and Speech: The Regulation of Hate Speech in a Democracy', *Wake Forest Law Review* 44 (2009): 497-532), but Waldron articulates the relationship between public order and hate speech more precisely than others who suggest such a relationship, his claim being that public order only obtains when certain public goods are secured, and that hate speech threatens public order by jeopardising one such good.

grounds of race, ethnicity, religion, gender, and in some cases sexual orientation.<sup>32</sup>

The idea of ‘assurance’, which crops up in the passage above, is an important element in Waldron’s account. In a well-ordered society the conspicuous features of the sensory environment – what is normally seen and heard in public spaces and forums – will support, or at the very least they will not undermine or conflict with, the state’s commitment to equality and justice.<sup>33</sup> Those commitments are by themselves an essential part of a society’s being well-ordered. But it is not enough, Waldron argues, that people’s statuses as equal members of society are formally enshrined in the legal system’s documents and operations. We also need to have a degree of security – including *felt* security – in these commitments obtaining in our society, and as Waldron observes, this need will be most keenly felt by members of historically oppressed groups.<sup>34</sup> What’s more, the assurances that people need (that they are positively owed, on Waldron’s view) are for the most part implicit and unspoken, and for that reason they are especially vulnerable to being undermined or subverted. Even if state Q’s labour laws are as a matter of fact formally and effectively egalitarian with respect to ethnicity or national origin, it still may only take one person’s public proclamation that “terrorist immigrants are stealing our jobs” to undermine, for many of Q’s workers, the confidence that they are secure and fully-entitled member of Q’s labour sector, and by

---

<sup>32</sup> p. 1599.

<sup>33</sup> Waldron is especially concerned with *visual* manifestations of identity-based hatred. With this emphasis he joins in a small but interesting sub-project in the free speech literature, of examining the ways in which non-verbal, symbolic representations convey hostile or oppressive meanings, and exploring the moral dangers that arise as a result (e.g. see David E. Whillock, ‘Symbolism in the Representation of Hate in Visual Discourse’ in Rita Kirk Whillock and David Slayden (eds.), *Hate Speech* (London: Sage, 1995); Christopher J. Newman and Peter Rackow, ‘Undesirable Posters and Dubious Symbols: Anglo-German Legal Solutions to the Display of Right-Wing Symbolism and Propaganda’, *Journal of Criminal Law* 75 (2011): 142-55). I’m not convinced, though, that Waldron’s argument needs to be wedded to an emphasis on the visual. His crucial idea is that of a *sensory environment*, and it seems plausible that this can be created as much by the ‘sound’ of hate speech as by the ‘look’ of it.

<sup>34</sup> p. 1634. The reference to ‘the well-ordered society’ and the idea of assurance as an element of social well-ordered-ness are both things Waldron borrows from John Rawls (*Political Liberalism* (New York: Columbia University Press, 1996)). Waldron does not attribute a position on anti-hate speech laws to Rawls, rather, he sees himself as adopting conceptual resources and normative themes in Rawls’ theory to formulate an argument in favour of such laws. It’s not clear, however, that the concept of assurance in Waldron has anything more than a nominal relation to Rawls’s claims about assurance. For Waldron, assurance is about “conveying to people a sense of security in their enjoyment of their most fundamental rights” (p. 1626). For Rawls, by contrast, the need for assurance derives from a game-theoretic problem. Just institutions rely upon people’s compliance, but if reasonable people are to comply with institutions that limit their liberty, they need to be assured that others will follow suit, so that in complying one is not making oneself easy prey to unscrupulous people (who may betray the institution for personal advantage). Certain social institutions, like the criminal law, may then be defended in view of their ability to provide such assurances. The provision of assurance *may* create a felt sense of security for people in their basic rights. But if and when it does, the moral import of that feeling derives from its tendency to promote institutional compliance, and the moral import of *that* derives in turn from the fact that compliance is needed to sustain just institutions. For Waldron, creating a felt sense of security seems like a moral end in its own right; that it may be, but this idea doesn’t evidently figure in Rawls’s thinking about assurance.

extension, of Q's society *per se*. Given the delicate and vulnerable nature of this good of assurance, Waldron argues that "the prime responsibility... that falls upon the ordinary citizen is to refrain from doing anything to undermine it or to make the furnishing of this assurance more laborious or more difficult", and on his view, this is the obligation that hate speech laws enforce.<sup>35</sup>

## 7. Status, welfare, and two rationales for assurance

Insofar as Waldron's argument is concerned with the bare *visibility* of hate speech, it aims to sidestep the empirical uncertainties that undermine many claims about hate speech's indirect harmfulness. Waldron is not telling a convoluted story about how hate speech is the root cause of economic inequalities, say, or discrimination in the courts. Rather, he is simply observing that in societies in which hate speech freely circulates – on the internet, on talk radio, and in books and pamphlets – the vilification of vulnerable social groups will be a familiar feature of the social environment. The observation seems uncontroversial. Waldron's principal contribution is to explain why it matters from a liberal legal perspective: (i) because providing assurance is an important part of securing people's dignitarian status, and (ii) because providing assurance to vulnerable groups is a delicate undertaking, which can be jeopardised by an environment 'disfigured by hate', even if only in isolated pockets. This seems plausible in its basic outline. Presumably part of the reason we ascribe moral significance to civic status is because we believe that social hierarchies are not just formally inequitable, but also painful or humiliating for subordinated parties. An appeal to the good of assurance is partly underwritten, then, by the insight that experiences of *de facto* subordination – whether they pervade every corner of a person's life, or just arise in certain settings and relationships – may still breed the kind of hurts and anxieties that come with subordination *per se*.

But how would hate speech be responsible for generating this sort of problem? There are two ways to explicate the details of Waldron's assurance-based argument for restricting hate speech, and both versions run into difficulty. At certain points Waldron seems to found his argument for restricting hate speech on the suggestion that anti-hate speech laws would be an effective way to protect the dignity-respecting elements of the existing social order, and to make our societies more respecting of dignity in the future. This reading of the argument is encouraged by passages like the following:

Even if a well-ordered society could dispense with laws prohibiting group defamation, it would be a mistake to infer from this that the societies we know must be prepared to dispense with those laws, as a necessary way of becoming well-ordered. Societies do not become well-ordered by magic. The

---

<sup>35</sup> p. 1630.

expressive and disciplinary work of law may be necessary as an ingredient in the change of heart within its racist citizens that a well-ordered society presupposes... [If] the good to be secured is a *public* good, a general and diffuse assurance to all the inhabitants of a society concerning the most basic elements of justice, then it is natural to think that the law would be involved.<sup>36</sup>

Given that a well-ordered society is one free from bigotry, this passage suggests, hate speech which manifests bigoted views will tend to impair our progress towards being socially well-ordered, and this is something the law can intervene in and counteract. Waldron acknowledges in his discussion that some liberals hold a contrary view: that well-meaning legal intervention in the ‘marketplace of ideas’ will in fact *decrease* the likelihood of the positive social changes that he envisages.<sup>37</sup> But he is sceptical about that view, and rightly so I think. Societies are not inherently geared towards their own edification. There are various reasons to favour an unfettered marketplace of ideas, but the Millian expectation that this will promote the progressive interests of humankind is, in many respects, just an article of faith. Instead of *laissez-faire* optimism, Waldron sees a case for using the law’s regimenting force to protect members of vulnerable social groups, by discouraging the most brazen forms of identity-oppressive expression.

However, while I think Waldron is right to doubt that racism will wither on its own in the ideological free market, the argument from assurance is robbed of its initial appeal when it is elaborated in this way, as an appeal to the potential long-term benefits of discouraging the most overt manifestations of identity-prejudice. Waldron initially seemed to be telling us that hate speech has an immediate, negative impact on people’s dignitary status. But when the argument is fleshed-out along these lines, it becomes an exercise in consequentialist speculation. Restrictions on hate speech *may* influence our societies away from identity-oppressive attitudes and practices over time, but they may not. Whether that is the case depends on how important a role hate speech plays in the complex causal systems – involving a mixture of institutional, material, economic, and discursive factors – out of which identity-based social hierarchies arise, and through which they are sustained. Hate speech might have an integral role in these systems, or it might have an entirely incidental role. Under a harm-prevention framework for assessing anti-hate speech law, this is the critical question. My point is that we do not make any progress on this question, simply by observing that assurance is an element of dignitarian status, and noting the possibility that restrictions on hate speech may be a good way to promote assurance in the long run. Under the first

---

<sup>36</sup> Waldron, *The Harm in Hate Speech* (2012), p. 81. In this instance I’ve quoted the most recent version of Waldron’s argument. The corresponding passage in the 2010 article is expressed in slightly less circum-spect language (see Waldron, ‘Dignity and Defamation’ (2010), p. 1623).

<sup>37</sup> The enormously influential ‘marketplace of ideas’ metaphor is usually attributed to US Supreme Court Justice Oliver Wendell Holmes, for discussion see Joseph Blocher, “Institutions in the Marketplace of Ideas”, *Duke Law Journal* 57 (2008): 821-89.

interpretation, these modest conclusions are all that Waldron's argument leads us to.

At other points, however, Waldron's argument from assurance seems to be appealing to the sort of considerations I sketched above, to do with the felt security of the hate speaker's targets, and its relation to their status as such. Even when person A understands that he is a fully-entitled member of his society as a matter of *de jure* right, the insecurity that comes with knowing he is despised by some of his neighbours may leave him humiliated, cowed, and unable to live the way that this status should make it possible for him to live. This reading of Waldron's argument is encouraged by passages like this:

We are talking about a display [in hate speech] that matters practically to individuals. It matters to them in their reliance on the principles of justice in the ordinary course of their lives, and in the security with which they enjoy that reliance. In a well-ordered society... People know that they can reasonably count on not being discriminated against, humiliated, or terrorized. They feel secure in the basic rights that justice defines; they can face social interactions without the elemental risks that interaction would involve if one could not count on others to act justly.<sup>38</sup>

In one respect, this version of the argument is on surer footing. There is no need for speculation here about how the toleration or restriction of hate speech may affect social order in the long run. The claim is that hate speech deprives its targets, here and now, of something they are owed, namely, the felt sense of security in their dignitarian status, which is partly constitutive of their having that status in the fullest and most complete sense.<sup>39</sup> The problem with this elaboration of Waldron's argument is that it paints a picture on which hate speech is not the actual source of the threat to people's assurance. What has gone awry, when assurance fails, is that people *cannot* reasonably count on not being discriminated against, humiliated, or terrorised. If hate speech was not backed by genuine prejudice or hostile intentions – if it was merely superficial, offensive abuse – then it seems doubtful that people's sense of assurance about their status would be seriously threatened. Hate speech gets its potency, in real life cases, from the fact that it *does* reflect a deeper, sinister current of identity-oppressive intention. But then it is surely this underlying state of affairs – the fact that there are people who

---

<sup>38</sup> p. 1627.

<sup>39</sup> Although, we might note, there is presumably still some need for evidence in establishing that *these* effects obtain, despite the fact that they are immediate, rather than long-term, consequences. In the absence of such evidence, Waldron might want to say that the undermining of assurance (in the sense of 'felt security') is an illocutionary act *constituted* by hate speech, rather than a perlocutionary *consequence* of hate speech. But that would significantly, and I think implausibly, alter his view. The claim is not that there is some hierarchical convention which is illocutionarily enacted by utterances of hate speech. The claim is that hate speech brings about insecure feelings for its targets, and that these feelings are constitutive of a certain kind of status-based harm. My point is just that the first part of this claim is a claim about the causal consequences of hate speech, and as such, it can be called into question with counter-evidence, or alternatively, made more persuasive with corroborating evidence.

genuinely feel the contempt they express in hate speech, and who would be prepared to act on it – which puts people in a position such that they cannot reasonably count on not being discriminated against, humiliated, or terrorised. Hate speech can of course be involved in the process through which these underlying states of affairs become known about. But for people in vulnerable, marginal, and historically-oppressed groups, the legal restriction of hate speech will not alleviate the assurance-eroding knowledge that there are others in their wider political community who feel contempt and hostility towards them. That knowledge can be conveyed and circulated in all sorts of subtle, indirect, culturally-codified ways – and of course, it can be (and where anti-hate speech laws are in effect, often is) powerfully conveyed in people’s preparedness to express their identity-based contempt even while faced with the threat or reality of prosecution.

In light of this strand of Waldron’s argument, it is useful to reflect on how we would think about someone who lived in a society in which she was subject to identity-based prejudice, but in which – due to sheltering, or extraordinary serendipity – she failed to realise that this was the case. One thought might be that someone ought to tell this person what she has not yet apprehended. Perhaps this will help her to stay on her guard against those who harbour maleficent intentions towards her. Perhaps there is a sense in which facing up to these terrible truths, about how one is seen by others, is itself partly constitutive of having one’s dignity in the richest and most robust sense.<sup>40</sup> In any case, even if we do not think one ought to relay the information that would disabuse this person of the dignitarian assurance they feel, there is a question about how we should understand the harm or loss that this person experiences when, having been confronted with hate speech, they come to see themselves as potential targets of humiliation and terror, and to see their status in the community as tenuous or threatened. When the hate speaker’s malevolent attitudes are brought out into the open, in a way that erodes people’s dignitarian assurance, it is the fact of the malevolence itself – as opposed to its being brought out into the open – that is fundamentally responsible for the erosion, and any harm constituted by it. If that is right, then there is a dispiriting conclusion awaiting us. When a group of people is held in contempt by others in their community, there is no way to secure the dignity of those belonging to the group in the fullest sense, short of eradicating the contempt itself. What we want is a social milieu in which people know that they will not be discriminated against, humiliated, or terrorised on account of their identity, because, as a matter of fact, it is *actually the case* they will not be discriminated against, humiliated, or terrorised on account of their identity. Granted, this aspiration brings us back to the first strand of Waldron’s argument. Anti-hate speech law may be one of the measures that we could employ to try to advance our societies along this trajectory. My response to Waldron, on that front, was that we need more evidence that the restriction of hate speech would in fact be a fitting

---

<sup>40</sup> This prospect is explored in Meir Dan-Cohen, "Harmful Thoughts", *Law and Philosophy* 18 (1999): 379-405, see pp. 402ff.

and effective measure to this end. As far as the second strand of the argument goes, my claim is that Waldron misidentifies the source of the problem. To the extent that hate speech undermines the dignitarian assurances that people are owed, it does so indirectly, by manifesting attitudes in view of which people may reasonably come to doubt the security of their status as fully-accepted and fully-entitled members of society.

## 8. Conclusion

In assessing the case for general legal restrictions on hate speech, our focus should be on whether hate speech indirectly harms its targets by contributing to the creation or perpetuation of identity-based social hierarchies, in which the targets are subordinated and disadvantaged. Under this approach it is useful and appropriate to characterise the harms of hate speech in dignitarian terms, our question being whether hate speech infringes upon the dignity of its targets and thus diminishes their status. But dignity-based analyses bring with them an ambiguity. There is one uncontroversial sense in which hate speech infringes against its targets' dignity, in that the hate speaker manifests his own refusal to ascribe his targets an equal status. However, for the purposes of justifying restrictions on hate speech, we need to see that hate speech infringes against its targets' dignity in the stronger sense that it diminishes their status in the wider social milieu that speaker and target inhabit. Heyman's argument from recognition purports to efface this distinction. Heyman says the hate speaker, in manifesting his refusal to ascribe dignity to his targets, thereby does diminish the targets' status in the wider milieu. I have argued that this ascribes too much significance to the actions and attitudes of isolated hate speakers. Waldron's argument has two strands. On one hand, he suggests that our legal treatment of hate speech will have a significant effect on how identity-based social hierarchies are sustained and replicated in the long run. The suggestion is not highly eccentric or implausible, but if we are going to mount a case for restricting hate speech on this basis, evidence must be added to conjecture. At other points, Waldron suggests that in order for people's dignitarian status to be maintained, people need to have a felt sense of assurance about the security of their place in the egalitarian social order. I have argued that this assurance cannot be secured through the legal restriction of hate speech, since the thing that undermines it is not hate speech itself, but the genuine identity-based contempt from which hate speech is motivated, and knowledge of which is available irrespective of whether hate speech circulates freely in the community.

In Heyman and Waldron's arguments we gain insights into the nature of identity-prejudicial social hierarchies and into the character of the harms that redound to people subordinated by such hierarchies. The mere fact that one is held in contempt by others on account of one's social identity is, in some respects, an impediment to one's holding one's place in a dignitarian social order in the fullest and

most robust sense. Being confronted by that contempt in the form of hate speech may add insult to injury, firstly, as Heyman emphasises, by making salient the fact that there are some who refuse to ascribe one the dignity one is owed as a fully-entitled member of society, and secondly, as Waldron emphasises, by further eroding a sense of felt security in one's dignitarian status. But it is the fact that there are attitudes of identity-based contempt in the community, rather than the manifestation of that contempt in hate speech, which is the fundamental source of the threat to people's dignity in this area. What we need, all parties will agree, is social reform such that these attitudes of identity-based contempt wane in their prevalence and potency. Securing the dignity of people in vulnerable social groups requires the cultivation of a social milieu in which identity-prejudicial speech is absent from public life, not because it is deterred by legal threat, but because its abject sentiments no longer command people's allegiance. Waldron thinks legal restrictions on hate speech have a role to play in driving such reforms, and the suggestion – given that it may found a justification for the restrictions on hate speech that operate in liberal democracies today – is one we must take seriously. To vindicate the suggestion, we need to show that hate speech is not just a peripheral element in the processes through which social hierarchies are perpetuated, nor an epiphenomenon created by identity-prejudicial social structures, but rather, that hate speech is a genuine contributor to the processes through which identity-prejudicial social hierarchies are reproduced. Hate speakers 'speak against' their targets' dignity, and they manifest the fact that the people they malign have not yet attained an equal dignitarian status, in the most complete sense. But these things do not establish that hate speech contributes to identity-based social hierarchies, in the way it would need to contribute, in order to substantiate the case for general legal restrictions on hate speech.<sup>41</sup>

---

<sup>41</sup> Thanks to Leslie Green, James Weinstein, Katherine Simpson, and two anonymous referees from *Law and Philosophy*, for valuable comments, criticisms, and corrections.