Towards a Legal Concept of Hatred: Democracy, Ontology, and the Limits of Deconstruction

The final version of this text appears in:

Hate, Politics and Law **(T. Brudholm & B. Johanssen, eds., forthcoming Oxford University Press, 2018)**

Abstract. Discrimination law focusses on material conduct. A legal concept of hatred, by contrast, focusses on attitudes, as manifest most notably through hate speech bans. Democracies by definition assign higher-law status to expression within public discourse. Such expression can, in principle, be legally curtailed only through a showing that it would likely cause some legally cognisable 'harm'. Defenders of bans have overtly or tacitly appealed to 'anti-Cartesian' phenomenological and socio-linguistic theories to challenge dominant norms which largely limit such harm to demonstrable material causation. Such notions of harm cannot, however, be reconciled with higher-law norms barring viewpoint-selective penalties on expression. Still, a democracy retains alternative means of combatting hateful attitudes, including formal and public educational policy, and codes of professional practice in the public and private sectors.

Keywords: antisemitism, critical legal theory, democracy, discrimination, hate speech, hatred, homophobia, Islamophobia, law & language, legal phenomenology, racism, sexism

Words: 9690

1. Introduction

Do we need a specifically legal concept of hatred? Must such a concept be distinguished from the concept, more familiar in law, of discrimination? The two concepts overlap, but are not identical. If they were, we would scarcely need to ponder hatred as a distinct problem for law. After all, we're already sitting on mountains of scholarship about discrimination.

[1] Some of the discussion in Section 6 is adapted from Section 5.3 of Heinze, E. *Hate Speech and Democratic Citizenship.* (Oxford: Oxford University Press, 2016).

We turn to the concept of hatred fearing that anti-discrimination policies may fall short. We worry that a narrow focus on discrete actions may neglect the underlying cultural attitudes fuelling those actions. Many observers insist that we must combat discrimination not only as a concrete type of action, but more fundamentally as attitudes embedded within culture. They warn against pat distinctions between conduct and attitudes, actions and ideas. They invite us to broaden our horizons beyond specifically discriminatory actions, so that we may target broader, discriminatory mind-sets.[2]

Psychological, sociological, and philosophical uncertainties admittedly haunt the concept of 'hatred', but that's not unusual for legal concepts within complex areas. Law must constantly address intricate social problems and develop vocabularies for them. In this essay, I shall certainly recognise various ways in which a concept of hatred can legitimately play that distinct role of targeting discriminatory attitudes beyond law's more familiar role of combatting materially discriminatory acts. I shall argue, however, that there are important limits to that role. Combatting discriminatory attitudes by penalising certain expressive acts – a task that the concept of hatred has increasingly been mobilised to perform – is in some respects an illegitimate task for the legal system of a fully-fledged democracy.

Free expression is generally deemed vital to democracy, notwithstanding disputes about its legitimate extent. To call expression vital means, in turn, that a democracy must show some serious countervailing interest if it is to impose legitimate penalties upon particular types of expression. Defenders of hate speech bans identify the need to combat discriminatory attitudes as one such countervailing interest. That claim, I argue in **Section 2**, presupposes evidence of a sufficiently strong causal link, such that hateful speech can be said to cause discriminatory attitudes in statistically demonstrable, hence more than haphazard ways. History has shown that link in many contexts, yet not in *longstanding, stable, and*

[2] See notes 5 - 6.

prosperous democracies (LSPDs), which develop fundamentally different social structures to foil the direct causation witnessed within non-LSPDs from hateful expression to systemically violent or discriminatory action. Advocates of bans respond by proposing alternative models of causation. Their challenge to a strictly material model of causation echoes longstanding doubts about what is commonly recognised as a Western, 'Cartesian' ontology, which simplistically distinguishes between physical and ideational spheres.

To be sure, distinctions between material actions and mental attitudes are by no means absolute in law. LSDPs, as I argue in **Sections 3 and 4**, do aim to shape public attitudes in legitimate ways – through primary education, public awareness campaigns, and the like. They can, moreover, appropriately penalise '*concomitant* attitudes', in other words, hateful expression accompanying independently criminal action, for the simple reason that motive-dependent penalties have long been integral to the legitimacy of criminal law. What remains questionable is whether the state can legitimate penalise '*essential* attitudes', that is, hateful expression in itself. In **Section 5**, I note an assumption that has been generally ignored by democratic theorists, elicited instead by democracy's critics, notably within post-phenomenological and post-Marxist schools. I accept that democracy's requirement of empirically demonstrable harm indeed presupposes a rightly challenged dualist, Cartesian ontology. I nonetheless conclude, in **Section 6**, that democracy does demand a limited yet certain commitment to a Cartesian ontology. To choose the most credible legal or political system is not to choose the one that would somehow 'overcome' ontological reductionism, mystically marrying politics to ontology, as many age-old idealisations, from ancient communitarianisms, to divine-right monarchies, to classical Marxism, to the theories of Heidegger or Schmitt, have often suggested. Any system for structuring a plurality of human interests, and therefore any legal system, always assumes ontological reductionisms.

Theories of non-empirical harm would demand that the state adopt assumptions incompatible with the ontology of democracy as a foundation for law.

2. Harm and causation

In democracies boasting strong human rights records[3], there scarcely remains serious controversy about the legitimacy of anti-discrimination law, at least as applied to core areas, such as primary education, basic goods and services, or routine workplace contexts. That largely settled core lies at one extreme end of a spectrum, where harm to victims is deemed patent. Such anti-discrimination law has never required a distinct concept of hatred in the familiar sense of 'great dislike or aversion' or 'intense ill will'.[4] Evidence of hostility can certainly enhance the case against a party accused of discrimination, but to *require* evidence of such sentiment would impede rather than advance anti-discrimination law. It would force claimants to demonstrate the indemonstrable – the emotional state of the accused. At that end of the spectrum, any evidence of hatred plays at most the subordinate role of further bolstering a discrimination claim, but cannot in itself decide the claim.

As we move away from that extreme pole, towards spheres in which the harms of discrimination are less certain, legal penalties become more controversial. We soon reach the opposite pole, where bans are imposed upon what is commonly called 'hate speech' – conduct exhibiting hostility towards particular social groups through solely verbal or visual means.[5] For proponents of hate speech bans, we must combat discrimination not only as material conduct, such as failure to hire qualified members of minority groups, but more

[3] See, e.g., EIU (Economist Intelligence Unit). *Democracy Index 2016.* (The Economist, London. 2017).
[4] McKechnie, J. L. (Ed.). *Webster's New Universal Unabridged Dictionary* 2nd ed. (Simon and Schuster, 1972), 831.
[5] See, e.g., Josende, L. *Liberté d'expression et démocratie : Réflexion sur un paradoxe.* (Brussels: Bruylant, 2010); Nieuwenhuis, A. *Over de grens van de vrijheid van meningsuiting: Theorie, rechtsvergelijking, discriminatie, pornografie* 3rd ed. (Nijmegen: Ars Aequi Libri, 2011); Pech, L. *La liberté d'expression et sa limitation: Les enseignements de l'expérience américaine au regard d'expériences européennes.* (Clermont-Ferrand: Les Presses Universitaires de la Faculté de Droit de Clermont-Ferrand, 2003); Noorloos, M. *Hate Speech Revisited: A Comparative and Historical Perspective on Hate Speech Law in the Netherlands and England & Wales.* (Mortsel, BE: Intersentia, 2012); Thiel, M. (Ed.). *Wehrhafte Demokratie: Beiträge über die Regelungen zum Schutze der freiheitlichen demokratischen Grundordnung.* (Tübingen: Mohr Siebeck, 2003).

fundamentally as a matrix of cultural attitudes. Sheer verbal or symbolic hostility, on that view, does not value-neutrally 'denote' or 'signify' hatred, but rather actively constructs – linguistically 'performs' – and in that sense systemically entrenches it.[6]

The harms of hate speech, for those who would ban it, are patent. Hate speech has fuelled atrocities within weak or fledgling democracies, such as the Weimar Republic, the former Yugoslavia, or Rwanda.[7] For today's leading democracies, however, we observe a decades-long failure on the part of governments, researchers, and activists to document statistically measurable links from hateful expression within general public discourse to such cognisable harms.[8] To be sure, those types of harms must be distinguished from the harms of hateful speech outside public discourse, in immediate, interpersonal harassment or 'fighting words' situations, for which infliction of emotional distress is in some cases demonstrable.[9]

I have elsewhere proposed the model of a historically recent political form, the *longstanding, stable and prosperous democracy* (LSPD), to describe a type of state that, far from persisting over centuries, has become recognisable only since the 1960s.[10] The LSPD

---

[6] Cortese, A. *Opposing Hate Speech.* (Westport, CT: Praeger, 2006); Delgado, R., & Stefancic, J.. *Must We Defend Nazis?: Hate Speech, Pornography, and the New First Amendment.* (New York: New York University Press, 1999); Delgado, R., & Stefancic, J. *Understanding Words That Wound.* (Boulder, CO: Westview, 2004); Gelber, K., & Mcnamara, L. " Evidencing the harms of hate speech." *Social Identities: Journal for the Study of Race, Nation and Culture., 22(3),* 2016; Langton, R. "Speech Acts and Unspeakable Acts." *Philosophy and Public Affairs 22*. 1993; Langton, R. "Beyond Belief: Pragmatics in Hate Speech and Pornography." In I. Maitra, & M. K. McGowan (Eds.), *Speech & Harm: Controversies over Free Speech.* (Oxford: Oxford University Press, 2012); Maitra, I., & McGowan, M. K. .*Speech & Harm: Controversies over Free Speech.*; Maitra, I., & McGowan, M. K. "Introduction and Overview.", In I. Maitra, & M. K. McGowan (Eds.), 1-23; Malik, M. "Extreme Speech and Liberalism." In I. Hare, & J. Weinstein (Eds.), *Extreme Speech and Democracy*. (Oxford: Oxford University Press, 2009), 96-120; Matsuda(et al. eds), M. *Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment.* (Boulder, CO: Westview Press, 1993); Waldron, J. *The Harm in Hate Speech.* (Boston, MA: Harvard University Press, 2012).

[7] See, e.g., Bemba, J. *Justice internationale et liberté d'expression: Les médias face aux crimes internationaux*. (Paris: L'Harmattan, 2008); Thiel, *Wehrhafte Demokratie.*; Tsesis, A. *Destructive Messages: How Hate Speech Paves the Way for Harmful Social Movements.* (New York: New York University Press, 2002).

[8] See, e.g., Meier, H. "Ob eine konkrete Gefahr besteht, ist belanglos": Kritik der Verbotsanträge gegen die NPD. In C. Leggewie, & H. Meier, *Verbot der NPD oder mit Rechtsradikalen leben?* (Frankfurt a.M.: Suhrkamp, 2002), 17, 22; Molnar, P., & Post, R. "Interview with Robert Post." In M. Herz, & P. Molnar (Eds.), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*. (Cambridge: Cambridge University Press 2012), 31; Neumann, V. "Feinderklärung gegen Rechts?: Versammlungsrecht zwischen Rechtsgüterschutz und Gesinnungssanktion." In C. Leggewie, & H. Meier, *Verbot der NPD oder mit Rechtsradikalen leben?* (Frankfurt a.M.: Suhrkamp, 2002), 363; Preuß, "U. K.. Die empfindsame Demokratie." In C. Leggewie, & H. Meier, *Verbot der NPD oder mit Rechtsradikalen leben?*, (Frankfurt a.M.: Suhrkamp, 2002), 110-11.

[9] See, e.g. Post, R. *Constitutional Domains: Democracy, Community, Management.* (Cambridge, MA: Harvard University Press, 1995), 114, 194; Weinstein, J. *Hate Speech, Pornography, and the Radical Attack on Free Speech Doctrine.* (Boulder, CO: Westview Press, 1999), 26–7, 72, 74, 76.

[10] Heinze, *Hate Speech,*69-78.

differs not only from non-democracies, but also from weaker democracies such as those in which hate speech has indeed unleashed crimes against humanity. Hateful expression by no means disappears from LSPDs. In sheer volume it may remain frequent, along with other, materially hence legally actionable manifestations of hatred. What, as an historical matter, distinguishes the top-performing[11] LSPDs is that they dispose of a range of legal, institutional, educational, and material resources to protect vulnerable groups or individuals from violence or discrimination in ways that do not require censorship of general public discourse.[12]

Advocates of hate speech bans nevertheless doubt that solely expressive action is harmless even in LSPDs. They challenge speech-conduct distinctions, which, in their view, reduce legally cognisable harms to material effects. The problem is not that liberal legal systems impose only standards of demonstrably material harm as the threshold criteria for legitimate limitations on individual freedoms. Local zoning laws, for example, often limit individuals' freedom to use property on entirely aesthetic grounds, without authorities needing to demonstrate that the regulations avoid harm in any conventional sense. Restrictions on citizens' expression, however, raise very different questions. Higher-law norms protecting expression within general public discourse furnish a necessary foundation for democracy in a way that ordinary rights in property or contract do not. Purely speculative allegations of harm caused by free expression cannot create grounds for limiting it without altogether destroying the meaning of a higher-order freedom.

---

[11] See, e.g., EUI, *Democracy Index.*

[12] See, e.g., Brettschneider, Corey. *When the State Speaks, What Should It Say?: How Democracies Can Protect Expression and Promote Equality*. (Princeton, NJ: Princeton University Press, 2012); EUI, *Democracy Index*.; Evans, C. "Religious Speech that Undermines Gender Equality." 368-73; Malik, *Extreme Speech and Democracy,* In I. Hare, & J. Weinstein (Eds.), 105-20; Molnar, P., & Post, R. "Interview with Robert Post." In M. Herz, & P. Molnar (Eds.), 22, 26, 32-34; Molnar and Strossen, interview, In M. Herz, & P. Molnar (Eds.), 392-93.

3. Punitive *versus* non-punitive approaches

LSPDs combat hateful attitudes largely through primary education, public awareness campaigns, and codes of best practice in employment and civic life.[13] Those constructive rather than punitive measures handily comply with democratic principles. We have already identified a legal concept of 'hatred' in an evidentiary role subordinate to that of anti-discrimination law; but here we see a more complementary or even interchangeable role between the concepts of discrimination and hatred. Where the state uses law to censure but not to censor – to express disapproval but without punishing those who express that which is disapproved – then 'pluralism', 'anti-discrimination', 'hatred', 'tolerance', and related concepts more easily blend with each other. Such policies do not demand the painstaking conceptual exegesis required of formally judicial situations, in which highly individual determinations of rights and liabilities are at issue.

That emphasis on 'constructive rather than punitive' measures is certainly not absolute. Under some circumstances public- or private-sector employees do rightly incur penalties for openly opposing pluralist values. Shop assistants or school teachers, for example, will face legitimate reprisals for using racist, sexist, or homophobic language in professional interactions. But those penalties sanction the offender only *as an employee*. Ordinary workplace situations have never been construed in law or in practice as full-fledged public forums admitting unbridled exchanges of ideas. Workplaces could scarcely function without constraints on expression.[14] The state may impose rules restricting disclosure of confidential information, forbidding false representations in business transactions, or conforming to other codes of dress and decorum which limit employees' expressive freedoms. The state may uphold dismissals for breaches even of the simplest courtesies,

---

[13] See note 12

[14] See, e.g., Jacobson, A., & Schlink, B. "Hate Speech and Self-Restraint.", 219-27; Molnar and Post, interview, In M. Herz, & P. Molnar (Eds.), 11-15; Weinstein, J. "Participatory Democracy as the Central Value of American Free Speech Doctrine." *Virginia Law Review, 97(3),* 2011, 491-95.

reaching into the minutiae of speech, such as compulsory recitations of 'Is there anything else I can help you with?', 'We hope you'll call again', and the like. Given both the pervasive speech regimes and the anti-discrimination policies that dominate employment situations, it would be arbitrary *not* to include penalties for hate speech.

In a word, democracies are certainly equipped to combat hateful attitudes. The thornier problem arises when a democracy penalises the citizen *as a citizen* by censoring expression within the general sphere of public discourse.[15] Western democracies recognise free expression within public discourse as a higher-law interest, conventionally through a human or civil right.[16] A freedom becomes higher-order if it can be abridged only through a showing that its exercise would pose a demonstrable risk of legally cognisable harm to other persons or to the state. Yet that criterion is by no means obvious. Law remains forever plagued by disputes about the existence or degree of harm.[17] As to free expression, risks posed by courtroom perjury, commercial fraud, or even excessive noise are admitted as grounds for curbing speech. Still, if a higher-order freedom is to have any meaning, it must require for its abridgment a showing of some risk that is more than speculative. That risk need be shown not for each and every exercise of the freedom, but across a class of activities. For example, commercial fraud or disclosure of nuclear secrets may be regulated because of sufficient risk posed in certain, sufficiently plausible instances, even if that risk is not present in every conceivable situation.

Some argue that the failure to link hateful public discourse to demonstrable harms lies with the genuine lack of any identifiable harm. Others insist that it lies with an ideologically

---

[15] On the character and contours of public discourse, see, e.g., Molnar and Post, interview, In M. Herz, & P. Molnar (Eds.), 12-17, 21; Post, *Constitutional Domains,* 7, 111, 134-77, 166 – 69, 194; Post, R. "Participatory Democracy and Free Speech." *Virginia Law Review, 97(3).* 2011, 483-84; Weinstein, *Hate Speech,* 168-76; Weinstein, *Participatory Democracy*, 495-96.

[16] Cf. Heinze, *Hate Speech,* 45-66 (maintaining that a higher-order freedom of expression must be deemed constitutive of democracy independent from concepts of individual rights).

[17] See, e.g., Heinze, E. Victimless Crimes. In R. Chadwick (Ed.), *Encyclopedia of Applied Ethics* 2nd ed. Vol. 4, (London: Elsevier/Academic Press, 2012); Kelly, P. "J.S. Mill on Liberty." In D. Boucher, & P. Kelly (Eds.), *Political Thinkers: From Socrates to the Present*. (Oxford: Oxford University Press, 2003).

biased notion that admits showings of harm only on material criteria. Those opposite interpretations create a dilemma. On the one hand, a higher-order freedom can legitimately be abridged only through a showing that its exercise would pose a demonstrable risk of harm to other persons or to the state; and to avoid the kind of open-ended state discretion that, by definition, would eviscerate a higher-order freedom, we correlate 'demonstrable' to empirical evidence. On the other hand, after decades of study, we altogether lack any such evidence within LSPDs.

One way out of that dilemma is to expand our concepts of risk by admitting non-empirical criteria of harm, sometimes identified as harm to human dignity[18], an interest that is certainly weighty albeit scarcely measurable. Far from overcoming the dilemma, however, that solution presents a new one. On the one hand, such a dignitarian impulse frees us from the shackles of empiricism. It vindicates the cardinal value of human dignity. On the other hand, as I shall argue, a state cannot 'free' itself from those 'shackles' without fundamentally impairing democracy's rule of law. Democracy, as opposed to outright majoritarianism[19], presupposes higher-order freedoms limited only by empirically demonstrable risks.

4. 'Essential' *versus* 'concomitant' attitudes

A hateful attitude is expressed by material means—the vibrations of vocal chords, ink drawn on paper, pixels on a screen. Hate speech bans do in that sense penalise material actions, just as the law punishes physical batteries or workplace discrimination. Unlike those latter acts, however, the material forms of expression are deemed harmful not inherently, but

[18] See, e.g., Barendt, Eric. *Freedom of speech*. 2nd ed. (Oxford: Oxford University Press, 2007), 31–4; Fiss, O. *The Irony of Free Speech.* (Cambridge, MA: Harvard University Press. 1996), 11–15; Grimm, D. "Freedom of Speech in a Globalized World.", 12; Heyman, S. J. "Hate Speech, Public Discourse, and the First Amendment.", In I. Hare, & J. Weinstein (Eds.), ,177; Morange, J. *La liberté d'expression.* Brussels: Émile Bruylant, 2009, 66; Parekh, B. "Is There a Case for Banning Hate Speech?", 43; Pech, *La liberté d'expression et sa limitation,* 149–62; Tillmanns, R. "Wehrhaftigkeit durch Werthaftigkeit--der ethische Grundkonsens als Existenzvoraussetzung des freiheitlichen Staates." In M. Thiel (Ed.), *Wehrhafte Demokratie: Beiträge über die Regelungen zum Schutze der freiheitlichen demokratischen Grundordnung*. Tübingen: Mohr Siebeck, 2003; Waldron, *The Harm in Hate Speech*, 39, 58, 60–1, 82–3, 86–7.

[19] See, e.g. Dworkin, R. "Foreword", in I. Hare, & J. Weinstein, vii.

solely through the hateful ideas they convey. Bodily attacks or unfair treatment at the workplace become particularly pernicious through hateful motives, but remain detrimental absent any link to hatred.

The law presents us, then, with two opposed paradigms for a legal concept of hatred. One paradigm we can call the model of *concomitant attitude*: the state takes hateful attitudes into account, but the act in question would cause detriment even without them. In addition to acts like physical batteries or discriminatory treatment in schools and workplaces, such acts can include distinctly psychological, and in that sense immaterial wrongs, such as infliction of emotional harm – which takes forms beyond hostility towards socially subordinated groups, but ordinarily involves immediate, individually targeted verbal assaults outside the sphere of general statements made within public discourse.[20]

The vibration of vocal chords, by contrast, or the setting of ink to paper in no way constitute harms in themselves. Under a second paradigm, which we can call the model of *essential attitude*, hate speech bans become controversial, at least from the standpoint of democratic theory, because no element of the unlawful act constitutes any kind of wrongdoing *except* the communication of the speaker's attitude. I am therefore proposing that contrast between the essential-attitude model and the concomitant-attitude model to replace the more customary distinction, namely, between *speech* as presumptively immune from regulation and *conduct* as a general object of regulation. That more conventional speech-conduct dualism fails, after all, to account for concomitant-attitude scenarios. (Note also that under neither the essential- nor the concomitant-attitude paradigms is it necessary that the attitude reflect the speaker's deeply felt views. Like any subjective feelings, such views can never conclusively be known. Elements of questionable irony or misjudged humour, for example, depend on particular fact patterns.)

[20] See, e.g. Post, *Constitutional Domains*, 119-34.

A comparison of the two paradigms leads to an inversion in the ways they respectively configure the relationship between the legal norm and the aberration. In cases of concomitant attitude, notably within criminal law, it would be aberrant to *exclude* evidence of hatred. Within post-Enlightenment criminal law, *mens rea* and motive remain integral to the very concept of wrongdoing. Under the essential-attitude model, however, we face the opposite state of affairs. Free expression within public discourse, limited only by demonstrable risks of at least moderate harm, is ordinarily the norm. Hate speech bans create an exception that requires some showing of harm insofar as free expression claims higher-law status.

Under the paradigm of concomitant attitude, hatred can be taken into account as part of a necessary *norm* for establishing some legally distinct wrongdoing. Under the paradigm of essential attitude, by contrast, attitude is taken into account as part of the *exception*—on a rationale that we must now examine—to an equally necessary norm that prohibits penalties solely on grounds that a repulsive idea is expressed. A supporter of hate speech bans might argue that they punish not the expression of ideas, but only of crude, 'visceral' invective of no greater ideational value than grunts and groans, a mere sub-speech. If so, it would seem, the state could proceed against the expression of hateful attitudes without censoring the expression of ideas.[21] As I have argued elsewhere, however, hate speech bans by definition assume the opposite. They cannot punish expression without discerning within it content that necessarily qualifies as ideational the moment they identify within it a hateful, or indeed any other, message.[22]

If the essential-attitude paradigm is to survive, then some broader concept of harm is required, beyond one that requires material causation, and which the concept of hatred might supply. The law must recognise a harm to dignity, construed as real yet not empirically

[21] See, e.g. Waldron, *The Harm in Hate Speech*, 335.
[22] Cf. Heinze, *Hate Speech*, 20-21, 108.

demonstrable, perhaps in the way that subjective experiences like 'self-esteem' or 'confusion' are real yet not – or not in any obvious or unqualified way – empirically demonstrable. It is insufficient to cite, as evidence of such a harm, wrongdoings like the aforementioned acts of infliction of emotional distress, or of reputational damage in defamation cases. As was noted, such discriminatory actions fall outside the types of general expression within public discourse of concern with respect to the regulation of individuals *as citizens*. We can already admit penalties for them under concomitant-attitude paradigms. Once we approach not 'speech' in general (a vast area indeed), but rather public discourse in particular as the distinct and essential medium of ongoing citizen participation within democracy[23], the burden of ascertaining any such dignitarian harm remains formidable.

5. The Cartesian framework

That broad concept of a dignitarian harm caused by general expression dovetails with challenges, familiar for well over a century, to classical Cartesianism as the dominant ontology of Western modernity. Writers committed to democracy have largely avoided examining democracy's ontological assumptions, leaving that task either to democracy's friendlier critics, like Walter Benjamin and Theodor Adorno, or to unfriendly critics like Martin Heidegger or Carl Schmitt. Theorists within standard democratic frameworks too readily overlook modern democracy's Cartesian assumptions.

Cartesianism admits two modes of being. Things exist either immaterially as ideas (*res cogitans*) or materially as corporeal substance (*res extensa*).[24] Descartes is often read in literal, ahistorical ways. His ontology proceeds through self-contained propositions claiming universal validity, as he contemplates his melting wax in a *chiaroscuro* of apolitical

[23] See, e.g. Heinze, *Hate Speech*, 45-55.
[24] Descartes, R. *Méditations métaphysiques [Meditationes de prima philosophia]*. (M. Beyssade, Trans.) (Paris: Livre de Poche, [1641] 1990), 55-63.

serenity.[25] It can be no accident, however, that Descartes pens that ontology as the horrors of the Thirty Years' War have been playing out.[26] The pivotal role of religious divisions in fuelling that pan-European conflict points the way towards rule-of-law notions that would, at least as a general matter, seek to steer the state towards empirically demonstrable harms and away from rules seeming to punish expressive acts solely as essential attitudes.[27] Once Descartes unhinges *res cogitans* from the effects of physical forces, the liberation of individual will, and therefore of a world of individual wills all able to act beyond material determinism, becomes inevitable.

The clearest sign of the deeply political character of Cartesianism lies in the manner in which the Cartesian ontology underpinning modern democracy has been examined most ardently by the critics of both, who indeed view them as two sides of the same – decadent – coin. A crucially Heideggerian critique tacitly influences today's critical theorists, who challenge requirements that material causation be demonstrated as a necessary condition for penalising hateful expression.[28] I shall nonetheless continue to argue that the modern democratic state cannot legitimately dismantle what is admittedly a troubling ontological dualism in order to justify the suppression of public discourse through dignitarian harms under the essential-attitude paradigm.

That limit on democratic redress by no means unmasks democracy as distinctly alienated from ontology, in the ways Heidegger suggests. To the contrary, there is no such thing as a legal system that faithfully recapitulates ontology. The complexities governing human situations entail questions of sheer judgment that can never in any unambiguous way follow from ontology. Non-democracies, such as medieval aristocracies, early modern monarchies, or modern autocracies, scarcely boast of self-evident ontologies – and scarcely

[25] Descartes, *Méditations métaphysiques,* 69.
[26] For a classic account see Schiller, F. *Geschichte des dreißigjährigen Krieges.* (Zürich: Manesse, [1790] 1998).
[27] See, e.g., Fennema, M. *Geldt de vrijheid van meningsuiting ook voor racisten?* (Amsterdam: Elsevier. 2009), 17; Preuß, *Verbot der NPD,* 11.
[28] See, e.g. Heidegger, M. *Sein und Zeit.* (Tübingen: Max Niemeyer, [1927] 1979), 89-101.

boast of some greater fidelity to their own professed or implied ontologies than what democracies can achieve. It is not news that professed ontologies of immanent unity and harmony[29] constantly translate into politics of arbitrary hierarchy and repression. The fact that Heidegger or Schmitt could embed their anti-Cartesianism within the most hateful regime in history surely amounts to more than a miscalculation.

From the perspective of victims' lived experiences, the encounter with hatred expressed within public discourse strikes more deeply than any mechanical model of abstract ratiocination can capture. Can the democratic *Rechtsstaat* legitimately intervene in that visceral experience through penalties imposed on views expressed within public discourse? The democratic conception of the rule of law places us in a bind. On the one hand, it restricts the concept of harm in ways challenged by anti-Cartesian understandings of the self and the social world. On the other hand, democracy—proceeding only imperfectly, like any legal regime, from ontology to normativity—presupposes an ontologically unsatisfying yet normatively compulsory limit defined indeed by Cartesian dualism.

Modern democracy is understood to be optimised through social and philosophical pluralism, beyond crude majoritarianism, opening its public spheres to competing worldviews. Contrary to critiques that would collapse active democracy into the 'empty shell' of liberalism, however, pluralism of worldviews does not equate with compulsory state neutrality towards them.[30] According to the age-old 'democratic paradox', democracy must host even outright opposition to itself.[31] It maintains legitimacy in fending off anti-pluralist attitudes only in conformity with its own constitutive norms of citizen participation. To proceed punitively against individuals *as* citizens, by punishing essential attitudes expressed

[29] See, e.g. Heinze, E. *The Concept of Injustice.* (London: Routledge, 2013), 50-78.
[30] Heinze, *The Concept of Injustice,* 111-16; Molnar and Post, interview, in M. Herz, & P. Molnar (Eds.), 26.
[31] Kelsen, H. *Vom Wesen und Wert der Demokratie.* (Tübingen: J.C.B. Mohr, 1920).

within spheres of public discourse, is to violate democracy's own ontology, which necessarily presupposes those spheres.

Every democratic constitution entails the possibility of its own abolition or amendment. Those processes make sense only as proceeding from public discourse. Through constitutional abolition or amendment provisions, democratic processes expressly present public discourse as something which those processes cannot legitimately constitute precisely because public discourse constitutes *them*. To block expression within public discourse on the essential-attitude paradigm is to attribute to democratic processes the legal authority to restrict the very activity which constitutes those processes in the first place.[32]

Those processes can legitimately regulate public discourse to prevent demonstrable risks, but not to impede the expression of the very attitudes—translatable and therefore comprehensible in public discourse only *as* ideas, however odious some ideas may be—for which public discourse as a political foundation exists. Democratic ontology of course equally presupposes some measure of respect among citizens, hence the legitimacy of punishing hate-motivated batteries and discrimination under the concomitant-attitude paradigm, and, more generally, of promoting pluralist civic education and policy. But whether, as dignitarian writers would have it, that ontology reaches so far as to authorise penalties on the expression of hateful attitudes within democratic public discourse is what now remains to be examined.

6. Phenomenological and socio-linguistic challenges

One may speak rightly or wrongly on countless issues of legislative policy, such as raising the top-rate of income tax or legalising hard drugs. To speak out against a subset of citizens as such, by contrast, seems to assail not merely this or that particular issue of policy,

[32] Cf. Heinze, *Hate Speech*, 6, 82, 90.

but rather the very possibility of pluralist democracy.[33] If a speaker, for example, calls communists 'scum', such speech may be crude, but, on their view, goes only to the targeted persons' ideas, not to their very existence as persons or as citizens—unlike referring to, say, an ethnic minority group in such terms. That dignitarian approach echoes theories of 'militant' (*wehrhafte* or *streitbare*) democracy, as pursued notably in post-World War II Germany. The theory deems hateful expression to be an attack on the dignity of individuals as citizens, beyond the scope of sheer ideas. For many German writers, hate speech is banned not only to protect vulnerable target groups but to protect democracy itself.[34]

By presenting hateful expression as inherently harmful, irrespective of measurable effects, the dignitarian stance aims to bypass the requirement of adducing empirically demonstrable harms as a justification for abridging higher-order expressive freedoms. Hate speakers, they maintain, cannot coherently challenge a law curtailing their own exercise of expression as citizens insofar as they would themselves deny the citizenship of others. An attitude denying others' citizenship can legitimately be penalised, then, through its sheer expression as an essential attitude.

Yet the criterion of 'denying others' citizenship' raises a problem. As part of the state's effort to combat hatred, it is generally applied to members of traditionally subordinated groups. Assume, however, that after reading about Soviet or Maoist atrocities, someone publicly exclaims that anyone expressing sympathy for those regimes 'is sub-human' or 'ought to be hanged'. The blurring of boundaries between disrespecting ideas and disrespecting persons works not for, but emphatically against the view that penalties on publicly expressed attitudes of hatred can plausibly be circumscribed to protect only vulnerable identity-groups – as witnessed also in endless debates about insult to religious

[33] Heyman, *Extreme Speech and Democracy,* in I. Hare, & J. Weinstein (Eds.); Waldron, *The Harm in Hate Speech*.
[34] See generally, e.g. Thiel, *Wehrhafte Demokratie.*

sensibilities.[35] That arbitrary border confirms the artificiality of the dignitarian assumption that hatred of communists attaches solely to their ideas while hatred of identity-groups attaches to persons as citizens. Dignitarianism does not overcome an artificial ontology, but only replaces one with another.

Some febrile theorising has taken place to confront the anomaly of a noxious cause which, within LSPDs, lacks any statistically confirmed link to a damaging effect. Such projects include phenomenological, socio-linguistic, or deconstructionist attempts to bypass questions about material causation by depicting hate speech as its *own* harmful effect. Theories of indirect harm have emerged that reject notions of language as a sheer medium of otherwise immaterial ideas. That is where the challenge to Cartesian dualism tacitly enters, notably since the 1980s, with the feminist approach of Catharine MacKinnon[36], the linguistic philosophy of Rae Langton[37], or the phenomenological approach of Charles Lawrence[38]. Langton examines hateful expression as *illocutionary* – not merely 'signifying' hatred but enacting and thereby socially constructing or creating it; and as *perlocutionary*, disseminating adverse psychological effects regardless of materially demonstrable impact.

Those theorists focus on the intangible yet pervasive power of language. Hateful expression, they argue, germinates cultures of intolerance, not always through discrete, causally traceable chains of events, so much as through gradual and cumulative effects. Its incorporeal character renders it not harmless, but *ipso facto* harmful: violence and discrimination arise not through clockwork mechanisms triggered in each case by discernible

[35] See, e.g., Cliteur , P., & Herrenberg , T. (Eds.). *The fall and rise of blasphemy law.* (Amsterdam: Amsterdam University Press, 2016); Cram, i. "The Danish Cartoons, Offensive Expression, and Democratic Legitimacy." In I. Hare, & J. Weinstein (Eds.); Hare, i. "Blasphemy and Incitement to Religious Hatred: Free Speech Dogma and Doctrine."
[36] MacKinnon, C. A. *Only Words.* (Cambridge, MA: Harvard University Press. 1996). Cf., e.g., Cortese, *Opposing Hate Speech* , 77-101.
[37] Langton, 1993. See also Langton, in I. Maitra, & M. K. McGowan (Eds.).
[38] Lawrence, C. "If He Hollers Let Him Go", in *Words that Wound,* 53-88.

speech acts, but through an entire culture to which those acts individually and cumulatively contribute.[39]

Theories of indirect causation end up not so much *wrong*, but, so to speak, *too right by half*. To designate the illocutionary and perlocutionary force of utterances as a legally cognisable harm in itself would reach so far into speech, including violence or offence in countless works of film, music, art, literature or colloquial speech, as to turn public discourse within LSPDs into the wholly subordinated object of a government operated license. Everyday sexist expressions like 'bitch' or 'cunt', or idioms tending to diffuse into public thought so as to denigrate disabled or otherwise different persons, such as 'idiot', 'moron', or 'fatso', make the point starkly. What kind of penalty shall we impose, for example, on someone uttering the phrase 'Dialogue of the deaf", thereby socially constructing the mental inferiority of a group far more vulnerable than many of our more protected groups? To view such examples as trivial, or de-fused through habit, is to refute the very theory of linguistic performativity that was supposed to provide an alternative model of harm, or of causation, in the first place. If a theory of indirect effect has any meaning, it is precisely the most normalised expressions which must be viewed as the most dangerous. Commenting on German television broadcasting in 2013, Alexander Kühn and Marcel Rosenbach write,

> German television viewers are surely monsters. Sunday upon Sunday, they demand a human sacrifice: shot, strangled, drowned, poisoned, pushed down the stairs, driven to suicide, frozen, or burned. A few weeks ago in [the police series] *Tatort* . . . the murder victim even had his fingernails ripped out. Nine million people watched.[40]

[39] Cf. Cortese, *Opposing Hate Speech*; Williams, S. H. "Democracy, Freedom of Speech, and Feminist Theory: A Response to Post and Weinstein." *Virginia Law Review, 97(3),* 2011, 603, 606-07.
[40] Kühn, A., & Rosenbach, M. (2013, July 8). *Das Grauen am Abend*. Accessed January 31, 2013, http://www.spiegel.de/spiegel/print/d-102241790.html.

Other weekdays, the authors note, include similar items, as part of a long term increase in violent shows. The authors nonetheless go on to observe that the programming 'reads like the counterpart of [recently] published criminal statistics, which show a nationwide decrease in murders and physical assaults'.[41] On a theory of indirect causation we would have to read such broadcasts not only as 'indicative', but as 'performative'. They would convey not, or not only, inert or neutral Saussaurian signifiers, indifferently transmitting one set of messages just as any other show would transmit other messages. Rather, the broadcasts would systemically construct violence as a social norm. The theories of indirect and diffuse effect sweep far too broadly since, if speech becomes action, hence a routine object of legal regulation through its performative character, then, for example, homophobes suddenly become correct in their insistence that pro-gay information, advocacy, or entertainment does not merely articulate, but rather materially 'imposes' a pro-gay viewpoint without the consent of some of those who are exposed to it.

Judith Butler challenges some of the critical schools, rejecting their view of a one-sided, wholly disempowering social construction of inferiority through hate speech. She argues that such provocations cannot be construed as generating purely passive victims. Rather, they serve as a mobilising force.[42] For example, calls for suppressing the publication of cartoon images of the Prophet Mohammed certainly followed the 2005 Danish controversy. A familiar claim was made, namely, that such expression serves to silence disempowered groups. But as Ian Cram observes,

> [A]llowing the offensive expression appeared to embolden a number of those who claimed to be offended. Public discourse in Europe was inundated by a range of Muslim perspectives and responses to the cartoons. Indeed, far from alienating Muslims from the state, or silencing them in public discourse, it

[41] [41] Ibid.
[42] Butler, J. *Excitable Speech: A Politics of the Performative.* (New York: Routledge, 1997).

> could be argued that these participants in public debate demonstrated a healthy commitment to the idea that they could shape the contours of public policy. In so doing, those protesting against the right to engage in speech offensive to particular religious communities ironically revealed the fortitude that can be demanded of all of us in a pluralistic liberal democracy.[43]

Advocates of bans often reply that those socio-linguistic theories, even if not justifying bans on all indirectly or diffusively harmful speech, nevertheless justify bans on the most extreme, crudest forms. That objection, however, far from supplementing the phenomenological, socio-linguistic or deconstructionist theories, renders them superfluous. The whole point of distinguishing indirect and diffuse effects of speech is to highlight attitudes conveyed amorphously, as a subtle and pervasive whole. If we are now to add independent criteria as to what qualifies as 'extreme and crude' from within that fluid mass, then the identification of indirect and diffuse effects loses its purpose as a distinct foundation for bans. Suddenly we are back at square one, simply making straightforward viewpoint-selective judgments. Theories of indirect causation cease to play any distinct role.

However cogently phenomenological or socio-linguistic theory may dissolve conventional formalisms, law by definition never gets beyond them. Law *is* formalism, even when uncodified or indeed unwritten. A world 'beyond' formalism is a world beyond law. Democratic, as opposed to purely majoritarian, normativity must necessarily retain an admittedly reductive Cartesian distinction between concomitant attitudes and essential attitudes. Speech-conduct distinctions are, after all, normally uncontroversial in democracies. We ordinarily accept that robbing a bank is punished, while publicly proclaiming, albeit in a purely offhand way (i.e., not as part of an already-active plan to do so), that one would like to

[43] Cram, *Extreme Speech and Democracy*, in I. Hare, & J. Weinstein (Eds.), 310.

rob a bank is not, even if such a declaration might encourage some people to rob banks. Advocacy of illegal activity is in a sense necessary for law reform: while cannabis possession may in some places be illegal, one must openly favour a right of possession if one is to agitate for legal change. Speech-conduct distinctions are distinctive of democracies. Undemocratic regimes routinely punish open advocacy of illegal conduct. Even in situations involving national security interests, earlier crimes of treason have largely narrowed to encompass only the disclosure of useful information, and no longer include purely general expressions of sympathy for the enemy. Hate speech bans, along with similar laws against 'incitement to hatred' or 'glorification of terrorism', represent exceptions. On the one hand, to the crimes of violence against vulnerable groups, they add penalties for advocacy of such crimes. On the other hand, such laws penalise even speech within public discourse to which no other such substantive crime corresponds, such as calling a particular group filthy, stupid, and the like.

The old 'sticks and stones' adage was never true, least of all for members of traditionally vilified groups. Even highly general invective within routine public discourse causes distress to targeted audiences. Causation of that type surely underscores the limits of ontological dualism, as hateful ideas come to be experienced through a visceral impact. The LSPD, a society avidly committed to pluralist democracy, can combat such attitudes on many fronts, starting with the earliest education and continuing into the workplace, the provision of goods and services, and participation in civic life. Insofar as democracy does not, however, merely allow free speech, but rather is itself constituted through the citizen's prerogative of expression within public discourse, that higher-order freedom cannot legitimately be abridged without a showing of some legally cognisable risk of harm beyond the fact that spectators may, albeit understandably, object to the exercise of that freedom. Our ongoing failure to identify any such risk within LSPDs has prompted observers to re-conceive the very ontology

of harm, which rightly supports those various types of pro-active state intervention. Such intervention, however, violates the ontological conditions of democracy itself—as nothing *but* the product of viewpoints exchanged within public discourse—when the state extends it to impose viewpoint-selective penalties within public discourse.

7. Conclusion

There can be no doubt that we need a legal concept of hatred distinct from the concept of discrimination. Discrimination law has always taken evidence of subjective animus into account, yet typically depends upon showings of material actions and omissions. The content and contours of human subjectivity admittedly remain indeterminate for legal analysis; and yet, however we may ultimately characterise hatred, it persists in many circumstances where evidence of such material actions and omissions is otherwise lacking. Hate speech bans therefore present a paradigm case for the view that the state must move beyond the confines of discrimination law. They are adopted on the view that the state must also regulate the formation and dissemination of attitudes.

Evidence of material causation from hateful attitudes uttered within public discourse to immediately discriminatory and violent actions remains patent for many non-LSPDs, most notoriously under circumstances such as those of the Weimar Republic, the former Yugoslavia, or Rwanda. In LSPDs, by contrast, stronger counter-currents to hate speech, both state-directed and within the population more generally, have largely broken those causal chains, as witnessed by the comprehensive lack of evidence of any such causation throughout decades of published research. Within LSPDs, the conflict between, on the one hand, an undoubted evil, and, on the other hand, absence of such legally cognisable harm traceable to it, has generated an ontology which designates hateful expression either as a *malum in se* or as a force proceeding through channels of indirect causation.

Even if we assume some degree of real plausibility for that ontology, that does not suffice to justify its translation into legally binding norms. It can never suffice, in eliciting a shortcoming of law, merely to point to its failure to reflect a plausible ontological position. No legal system can straightforwardly recapitulate ontology in any such way. Democracy's constitutive commitment to the citizen's higher-law freedom of expression within public discourse means that viewpoint-selective penalties on such expression can only be imposed on grounds of risks more immediately and more specifically traceable to that expression than has ever been demonstrated within LSPD frameworks. That necessary outer limit to state action against hatred in no way, however, assumes a 'value neutral' state. To the contrary, the diffusion of direct material causation from hatred within public discourse to legally cognisable harms has resulted largely from the success of LSPDs in combatting hateful attitudes in many other areas, in ways indeed surely more beneficial to targeted groups. Those alternatives to viewpoint-selective penalties imposed within public discourse emphasise the promotion of pluralist democracy in education, at the workplace, in the mass media, and in civil society.