State Inspection

Jeremy Waldron[*]

How should we think about the harm caused by hate speech? Does Oliver Wendell Holmes's "clear and present danger" test provide an appropriate framework for evaluating such harms? I think not, and I would like to begin with an analogy.

A week ago, I received an email advising me that the required annual New York State Inspection for my aged Subaru must take place by July 31, 2020. The notice commanded me to take the car to a licensed inspection station to have it checked for, among other things, a properly functioning emissions-control system. This is a nuisance and it costs money, even if the car passes the inspection. And if it does not pass, modifications and repairs to the emissions system can be quite expensive. Still, the test and the remedies for defects are required by law. They are required because nothing but ensuring that all vehicles have good working emissions systems can abate the danger posed by the accumulation of pollutants from millions of individual vehicles.

New York State is obligated to require all this under the auspices of the federal Clean Air Act.[1] The Environmental Protection Agency (EPA), which administers that legislation, says the following in one of its pamphlets on emissions:

> Emissions from an individual car are generally low, relative to the smokestack image many people associate with air pollution. But in numerous cities across the country, the personal automobile is the single greatest polluter, as emissions from millions of vehicles on the road add up. Driving a private car is probably a typical citizen's most "polluting" daily activity.[2]

[*]University Professor, New York University School of Law.

[1] Emissions testing rules in New York are set out in Regulations of the Commissioner of Motor Vehicles (Part 79, Motor Vehicle Inspection), N.Y. COMP. CODES R. & REGS. tit.15, § 79 (2020), as authorized and required under section 301(d) of the Vehicle and Traffic Law, N.Y. VEH. & TRAF. LAW § 301(d) (Consol. 2020). That is in turn required under Title II of the federal Clean Air Act.

[2] U.S. ENVTL. PROT. AGENCY, AUTOMOBILE EMISSIONS: AN OVERVIEW 1 (1994), https://www.wisconsinvip.org/WivipPublic/PDF/Downloads/2_2_4.pdf.

So here is my question: can I legitimately complain about the State Inspection requirement because—as the EPA acknowledges—the level of pollution caused by my car, or any individual car, is rather low even if its emissions system is broken? In and of itself, the operation of my Subaru poses no significant emissions hazard to anyone. There must be thousands of such cars with non-working emissions systems before there is discernable danger. So would I be justified in calling for the repeal of the relevant statutes and regulations because of the unwarranted burden they impose person-by-person on individual car-owners?

We all know that such a response would be silly. As a society, we collectively face dangers of air quality deterioration, carbon monoxide poisoning, and lead toxicity, not to mention the greenhouse gas effects of CO_2. The fact that these dangers are not singularly caused by the isolated acts of individuals considered on their own is not a reason to refrain from abating them. There is no way to deal with these dangers except by restrictions on drivers like me. The danger is real, each car contributes, and emissions from large numbers of exhaust pipes add up. Nothing but restrictions on thousands of cars like mine—requiring testing and working emissions-control systems—offers any prospect of averting the danger.

The temptation is to think that since regulation bears down upon the behavior of individuals one-by-one, an individual should be regulated only with respect to the harm caused specifically by his or her actions. It is unjust, we might have said, to hold A liable for something done by B. But if the similar actions of A and B and C (and thousands of others) add up to a hazard, H, that impacts adversely on public health generally, there may be no way of abating H except by regulating A's actions considered in light of B's and C's, and B's actions considered in light of A's and C's, and C's actions considered in light of A's and B's, and so on. An individualized focus on A's actions alone, without reference to the accumulation of its effects with the effects of the actions of B and C, will leave H unabated and many people considerably worse off as a result.

Even moral philosophers now understand this. The 1984 publication of Derek Parfit's *Reasons and Persons* marked a turning point in consequentialist analysis.[3] Parfit named Chapter Three "Five Mistakes in Moral Mathematics." Among the mistakes he diagnosed were "[i]gnoring the [e]ffects of [s]ets of [a]ctions" and "[i]gnoring [s]mall or [i]mperceptible [e]ffects." As to the first, he wrote that "[i]t is

[3] DEREK PARFIT, REASONS AND PERSONS (1984).

natural to assume . . . [that] [i]f some act is right or wrong *because of its effects*, the only relevant effects are the effects of this particular act."[4] But he invited us to consider cases, familiar to theorists of criminal law and tort law, where several people acting independently have an impact on a single individual, resulting in an injury whose severity cannot be attributed to any one of them alone.[5] And as to the second, he posited a case in which an individual is tortured by having the electric current on an apparatus turned up gradually by imperceptible increments, thousands of which cumulate into the infliction of agonizing pain.[6] A single torturer turns the increment switch again and again until the victim is screaming in pain. Can the torturer defend himself against the charge of torture by saying that nothing he did in any one of his actions had a discernible effect upon the victim? Or put the two cases together: Each of a thousand torturers turns his switch once to increase the current by a tiny amount. If they understand the situation, are they not all of them complicit in a set of actions that has dire consequences? Or can they take refuge in absence of any direct and clear causal line extending from what each does individually to what the victim eventually suffers? Parfit's hypotheticals are curious in their substance,[7] but the moral is clear: When we consider the consequential significance of what we do, we must take into account the actions of many individuals. The consequences of these collective actions can result in a harm, the magnitude and importance of which is belied by focusing on each person's behavior, considered under the auspices of older and cruder models of causation.

In the pollution cases, the accumulation of overall harm takes place over more than one dimension. The effects that H comprises may be an accumulation across thousands of individual actions occurring simultaneously. Toxins in the air breathed by children in an urban playground may be the result of thousands of vehicles operating in the vicinity on a given morning. Or H may build up slowly over time: the toxins present in a town's water supply may be the result of years of pollution—again at a level that is barely perceptible considered as an individual effect—by vehicles driving at or near that town. Once the responsible authorities become aware of the danger and its causes, they can intervene with regulation and controls to make it less likely that the

[4] *Id*. at 70, 75.

[5] *See* Summers v. Tice, 199 P.2d 1, 3 (Cal. 1948) (establishing doctrine of alternative liability).

[6] *See* PARFIT, *supra* note 3, at 80–82.

[7] For doubts about some of Parfit's hypotheticals, see Kristin Shrader-Frechette, *Parfit and Mistakes in Moral Mathematics*, 98 ETHICS 50 (1987).

harm will accrue. Such regulation prevents the accumulation of the large number of tiny indiscernible effects that would otherwise grow together to constitute a deadly environmental effect.

This strategy, however, is possible only if the authorities refuse to limit their justifying restrictions to imminent and directly traceable harm, in other words, harm that accrues immediately from each individual's actions. As Cass Sunstein points out, "regulators often act without the slightest hesitation even though the benefits of their action will not be immediate; indeed, such benefits might occur many years in the future (as in the case of climate change)."[8] They have to do so because waiting until the harm was imminent would make abatement impossible.

Maybe Sunstein is too optimistic about regulators' attitudes to harms of this kind. Many environmentalists are far from convinced that the politics of regulation can deal with what they call "slow threats," i.e., situations "where small, hardly noticeable changes add up over time to produce large impacts."[9] The time scale may be decades or generations. It is not easy to factor the pacing of such threats into the frameworks of modern politics. Their accrual involves "time periods that historians regularly deal with, but that stretch out beyond the time frame in which governments make budgets or do strategic planning."[10] Such threats

> may be mentioned occasionally in newscasts or op-eds when a major research report is published or some disaster occurs, but they seldom reach the level of sustained visibility and concern they deserve. Without that awareness and sense of alarm, the problems are likely to continue worsening until their impacts become severe and obvious, stressing our ability to respond, or, in the worst cases, passing tipping points where no amount of effort can prevent catastrophic impacts.[11]

*

The "Slow Threats" literature shows that progress in regulatory approach is still patchy. It seems to be particularly problematic in regulating social harm—especially pervasive social harm that accrues from the accumulation of large numbers of relatively small-scale phenomena, such as individual speech acts, interactions, and aggressions.

---

[8] Cass R. Sunstein, *Does the Clear and Present Danger Test Survive Cost-Benefit Analysis?*, 104 CORNELL L. REV. 1775, 1777 (2019).

[9] Robert L. Olson & David Rejeski, *Slow Threats and Environmental Policy*, 48 ENVTL. L. REP. NEWS & ANALYSIS 10116, 10116–17 (2018).

[10] *Id.* at 10117.

[11] *Id.* at 10116–17.

One area where accumulating harm has been taken seriously concerns the accumulation of small-scale but visible damage to buildings and vehicles and its effects on the overall morale of urban neighborhoods. The "broken windows" theory was put forward by George Kelling and James Q. Wilson in a 1982 article,[12] in which they argued that visible signs on the streets of lack of repair—things like a broken window, a pile of uncollected garbage, damaged furniture, or an abandoned car—were likely to attract further disorder and crime because they signified to anyone interested that social controls were weak or attenuated at that locale.[13] No single item would lead to this effect, but an accumulation, even a slow accumulation of such damage over time, would transform the atmosphere on that block or in that neighborhood.

Now some commentators have made a direct connection between the broken windows theory and the damage wrought by hate speech. Gregory Rodriguez says that "[h]ate speech is a form of vandalism. It defaces the environment, and like a broken window, if left untended, signals to other hoodlums that the coast is clear to do more damage."[14] Maybe this is how some of the harm from hate speech comes about. And sometimes it may be intentional. Frank Collin, the leader of the Nazis who marched through Skokie in a famous incident from the 1970s said: "We want to reach the good people—get the fierce anti-Semites who have to live among the Jews to come out of the woodwork and stand up for themselves."[15] The idea was that timid anti-Semites would feel safer expressing themselves publicly in a social atmosphere already vandalized by Collin's neo-Nazi followers parading with foul and provocative signs ("Hitler should have finished the job," etc.). But there are also broader modes of the accumulation of social harm that are more akin to the slow-acting poison hypothesis that regulators use in understanding environmental threats.

As I have indicated, developing an appropriate regulatory response to the harm done by hate speech is made difficult by the iconic status of the "clear and present danger" test in First Amendment law and politics. I suspect the currency of Oliver Wendell Holmes's formulation indicates that some of us are still in the grip of primitive thinking in some of what

---

[12] George L. Kelling & James Q. Wilson, *Broken Windows*, ATLANTIC MONTHLY, Mar. 1982, https://www.theatlantic.com/magazine/archive/1982/03/broken-windows/304465/.

[13] This language is adapted from my article, Jeremy Waldron, *Homelessness and Community*, 50 U. TORONTO L.J. 371, 386 (2000).

[14] Gregory Rodriguez, *Vandalized by Speech*, L.A. TIMES, Mar. 26, 2012, at A13.

[15] PHILIPPA STRUM, WHEN THE NAZIS CAME TO SKOKIE 15 (1999); *see* JEREMY WALDRON, HARM IN HATE SPEECH 95 (2012).

we say about hate speech and the harm it causes. According to Holmes, it is permissible to "punish speech that produces . . . a clear and imminent danger that it will bring about forthwith certain substantive evils that the United States constitutionally may seek to prevent. . . . It is only the present danger of immediate evil or an intent to bring it about that warrants Congress in setting a limit to the expression of opinion . . . ."[16] The reiteration of these requirements of fast-approaching danger—"clear and imminent," danger "that it will bring about forthwith," "the present danger of immediate evil"—raises doubts about the serviceability of Holmes's doctrine in addressing slow threats and the accumulation of harm. As I said in my book *The Harm in Hate Speech*,

> Imagine if we took that attitude toward environmental harms—toward automobile emissions, for example. Suppose we said that unless someone can show that *my automobile* causes lead poisoning with direct detriment and imminent harm to the health of assignable individuals, I shouldn't be required to fit an emission-control device to my car's exhaust pipe. It would be irresponsible to reason in that way with regard to environmental regulation; instead we figure that the tiny impacts of millions of actions—each apparently inconsiderable in itself—can produce a large-scale toxic effect that, even at the mass level, operates insidiously as a sort of slow-acting poison, and that regulations have to be aimed at individual actions with that scale and that pace of causation in mind.[17]

Why then is it sensible to take that approach—looking only to macro-harms and their immediate causation—when we are considering social harm? As Leslie Kendrick observes, "[t]he clear and present danger standard limits itself to acute situations, where particular expression is highly likely to have an immediate, demonstrable impact."[18] It would not cover "speech that erodes democratic ideals gradually—speech that promotes anti-Semitism and racial hatred," for example.[19] "Repeated, corrosive exposure causes harm by

[16] Abrams v. United States, 250 U.S. 616, 627–28 (1919) (Holmes, J., dissenting). The actual phrase "clear and present danger" came not from *Abrams* but from the slightly later case of *Schenck v. United States*, 249 U.S. 47, 52 (1919). "The question in every case is whether the words used are used in such circumstances and are of such a nature as to create a clear and present danger that they will bring about the substantive evils that the United States Congress has a right to prevent. It is a question of proximity and degree." *Id.*

[17] WALDRON, *supra* note 15 at 97 (emphasis in original).

[18] Leslie Kendrick, *On "Clear and Present Danger"*, 94 NOTRE DAME L. REV. 1653, 1665–66 (2019).

[19] *Id.* at 1666.

accumulation," says Kendrick, but "[a] literal reading of Holmes's . . . opinions would suggest that such harms are not the type of emergency that justifies suppression."[20]

How does all this apply to hate speech? Doctrinally, Kendrick's difficulty with the clear and present danger test is difficult to see because those who favor restricting hate speech sometimes argue as though such speech is per se valueless. On that argument, the "clear and present danger" test would not apply at all, since hate speech would be in the same category as obscenity or constitutionally unprotected defamation.[21] And indeed the Supreme Court in *Beauharnais v. Illinois* suggested that hate speech might be put in this category by being considered as group defamation.[22] This certainly provides an alternative to the clear and present danger test, but the alternative does not require us to vary our mode of consequential analysis to include non-imminent threats. In fact, defenders of hate speech regulation have sometimes been willing to pursue both lines of argument. Thus, on the one hand, the racist pamphlet at issue in *Beauharnais* was classified as a group libel. On the other hand, Justice Frankfurter was also willing to approach the issue of consequential harm by making a powerful argument about social atmospherics—the danger of "exacerbat[ing] tension between races."[23]

> Illinois did not have to . . . await the tragic experience of the last three decades to conclude that wilful purveyors of falsehood concerning racial and religious groups promote strife and tend powerfully to obstruct the manifold adjustments required for free, ordered life in a metropolitan, polyglot community.[24]

The either/or here was not valueless speech versus a showing of clear and present danger; it was valueless speech versus a showing of harm, which might include *either* clear and present danger *or* slow-acting harm to the social environment.

---

[20] *Id*. at 1667.

[21] *See* Tom Hentoff, *Speech, Harm, and Self-Government: Understanding the Ambit of the Clear and Present Danger Test*, 91 COLUM. L. REV. 1453, 1462–63 (1991).

[22] *See* Beauharnais v. Illinois, 343 U.S. 250, 266 (1952) ("Libelous utterances not being within the area of constitutionally protected speech, it is unnecessary, either for us or for the State courts, to consider the issues behind the phrase 'clear and present danger.' Certainly no one would contend that obscene speech, for example, may be punished only upon a showing of such circumstances. Libel . . . is in the same class.").

[23] *Id.* at 259.

[24] *Id.* at 258–59.

This consequential analysis is certainly what the best-drafted hate speech statutes around the world seem to be looking to capture. The United Kingdom's Public Order Act looks at the speaker's intention "to stir up racial hatred."[25] The German Penal Code and the Canadian Penal Code are concerned with the incitement of hatred, the New Zealand Human Rights Act punishes words "likely to excite hostility," and so on.[26] All of these measures seem to be concerned with the effect of hate speech in poisoning the atmosphere in which diverse groups are expected to live together.

I suppose one could parse such statutes in a spirit of strict individualism: one could interpret them as prohibiting one person, X, from stirring up hatred in the mind of another person, Y, against a third person, Z. But the hatred that the legislators are concerned about seems to diffuse community-wide hatred against members of a large (albeit minority) group defined by race, religion, or national origin rather than hatred focused on particular individuals. The hatred itself is mentioned as something abstract, spread directly and indirectly among members of the audience to whom the message of hate is delivered. And it makes sense to understand the offense as a person doing something—in circumstances in which others are likely to be doing it too—so that what

[25] Public Order Act 1986, c. 64, § 18 (1) (Eng.) ("A person who uses threatening, abusive or insulting words or behaviour, or displays any written material which is threatening, abusive or insulting, is guilty of an offence if—(a) *he intends thereby to stir up racial hatred*, or (b) having regard to all the circumstances *racial hatred is likely to be stirred up thereby.*") (emphasis added).

[26] STRAFGESETZBUCH [StGB] [PENAL CODE], § 130(1), *translation at*, https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html (Ger.) ("Whosoever, in a manner that is capable of disturbing the public peace 1. *incites hatred against segments of the population* or calls for violent or arbitrary measures against them; or 2. assaults the human dignity of others by insulting, maliciously maligning, or defaming segments of the population, shall be punished with imprisonment from three months to five years." (emphasis added)); Criminal Code, R.S.C. 1985, c C-46 § 319(1) (Can.) ("Everyone who, by communicating statements in any public place, *incites hatred against any identifiable group* where such incitement is likely to lead to a breach of the peace is guilty of (a) an indictable offence and is liable to imprisonment for a term not exceeding two years" (emphasis added)); Human Rights Act 1993, s 61(1) (N.Z.) ("It shall be unlawful for any person—(a) to publish or distribute written matter which is threatening, abusive, or insulting, or to broadcast by means of radio or other electronic communication words which are threatening, abusive, or insulting; or (b) to use in any public place . . . or at any meeting to which the public are invited or have access, words which are threatening, abusive, or insulting; or (c) to use in any place words which are threatening, abusive, or insulting if the person using the words knew or ought to have known that the words were reasonably likely to be published in a newspaper, magazine, or periodical or broadcast by means of radio or television,—being matter or *words likely to excite hostility against or bring into contempt any group of persons in or who may be coming to New Zealand on the ground of the colour, race, or ethnic or national origins of that group of persons*" (emphasis added)).

has to be considered from the legislators' point of view is the effect of this one incident of hate speech alongside lots of other such incidents in society at large. Of course, one act of individual hate speech may fizzle out without any larger social effect, like one of the ratty little pamphlets that Holmes alluded to in *Abrams*.[27] But before we can dismiss it in that way, we must be sure that its apparent harmlessness is not an illusion born of the assumption that long-term social degradation does not matter.

The long-term threat envisaged in this legislation is the stirring up and spreading of hatred and contempt, of hostility against and between various groups in society. In any diverse society, inter-communal hostility and the undermining of social peace is one of the things most to be dreaded. Once it comes into existence, it is very hard to roll back. One need only think of Hindu-Muslim antipathy in India, for example, or, for much of the twentieth century, hatred between Nationalist (Catholic) and Loyalist (Protestant) communities in Northern Ireland. Or consider the effects of immigration. The New Zealand provision mentioned earlier makes an explicit connection to the social tensions that might arise out of an influx of immigrants; it talks of "words likely to excite hostility against or bring into contempt any group of persons in *or who may be coming to* New Zealand."[28] The authorities have to pay attention to such possibilities. Those responsible for the well-being of a society cannot view this with indifference or, given the mode of its causation, postpone dealing with it until the onset of its worst effects is imminent.

One can also look at the issue of social atmospherics in a more positive light. Think of Justice Frankfurter's modest formulation in *Beauharnais*: "the manifold adjustments required for free, ordered life in a metropolitan, polyglot community."[29] These adjustments define for us the deep ecology of respect, tolerance, and inclusion in our society. We are diverse in our ethnicities, our race, and our religion. And we are embarked on a grand experiment of living and working together within and despite these differences. Each group must accept that society is not just for them, but it is for them too, along with all the others. And each person, each member of each group, should be able to go about his or her business with the assurance that there will be no need to face hostility, violence, discrimination, or exclusion by others. When this assurance is conveyed effectively, it is hardly noticeable; it is something on which everyone can rely, we hope, like relying on the cleanness of the

[27] Abrams v. United States, 250 U.S. 616, 628 (1919) (Holmes, J., dissenting).

[28] Human Rights Act 1993, s 61(1) (N.Z.) (emphasis added).

[29] Beauharnais v. Illinois, 343 U.S. 250, 259 (1952).

air they breathe or the quality of the water they drink from a fountain. This sense of security in the space we all inhabit is a public good, and in a decent society, it is something that we all contribute to and help sustain in an instinctive and almost unnoticeable way. Hate speech undermines this public good, or it makes the task of sustaining it much more difficult than it would otherwise be.[30]

This is best stated by Bhikhu Parekh, a British political theorist and member of the House of Lords:

> Hate speech . . . lowers the tone of public debate, coarsens the society's moral sensibility, and weakens the culture of mutual respect that lies at the heart of a good society. . . . It creates barriers of mistrust and hostility between individuals and groups, plants fears, obstructs normal relations between them, and in general exercises a corrosive influence on the conduct of collective life.[31]

Ultimately the impact is on individuals, as public expressions of hatred change the character of the environment in which they must live and work. In *R. v. Keegstra* (1990), the Canadian Chief Justice Brian Dickson said this about the effect that public expressions of hatred may have on people's lives:

> The derision, hostility and abuse encouraged by hate propaganda . . . have a severely negative impact on the individual's sense of self-worth and acceptance. This impact may cause target group members to take drastic measures in reaction, perhaps avoiding activities which bring them into contact with non-group members or adopting attitudes and postures directed towards blending in with the majority. Such consequences bear heavily in a nation that prides itself on tolerance and the fostering of human dignity through, among other things, respect for the many racial, religious and cultural groups in our society.[32]

A hateful atmosphere stigmatizes members of vulnerable groups, contributing, as Parekh puts it, to "a climate in which, over time, some groups come to be demonized and their discriminatory treatment is accepted as normal."[33] Parekh points out the obtuseness of working with a Holmesian standard in this area. "[H]arm," he says,

> is not easy to define, identify and prove. A relentless decline in the society's climate of mutual respect is a case of serious harm, but it does not point to specific individuals. . . .[.[A]n

[30] This paragraph is adapted from WALDRON, *supra* note 15, at 4.

[31] Bhikhu Parekh, *Limits of Free Speech*, 45 PHILOSOPHIA 931, 932 (2017).

[32] R. v. Keegstra, [1990] 3 S.C.R 697 (Can.).

[33] Parekh, *supra* note 31, at 933.

> isolated case of hate speech can be taken in one's stride. But over time such remark [sic] create a climate that might be deeply offensive and unacceptable. [34]

"Vicious and widespread hatred of a group does not spring up overnight."[35] This is clearly danger, but it is not "present" danger in Holmes's sense. Like any of the slow threats we described earlier, it "builds up slowly, through isolated utterances and actions, each perhaps trivial individually but all cumulatively capable of coarsening the community's sensibility."[36]

I think this is what we ought to be addressing in our thinking about hate speech. It is true that in *Abrams v. United States*, Holmes gave voice to a rightly celebrated dissent. But he did not leave us with the doctrines we needed in order to think sensibly about this kind of harm generated in this kind of way.

*

I want to end with some thoughts about the acceptance or lack of acceptance of this kind of argument. In debates about hate speech—at least in the United States—whenever I make a case for paying attention to slow-building atmospheric harm of the sort described here, I am confronted with immediate skepticism. Audiences flatly insist that we must stick with something like the "clear and present danger" approach and they will barely countenance any other approach to the causation of harm. And so, I wonder, what explains this insistent reliance on Holmes's standard when its use would make little sense in other areas of regulation? Why this blind loyalty to what seems like a crude and outdated understanding of the way harm accrues and the way it may be regulated?

The sense I get is that people fear some *trick*; they suspect this is not real causation at all, just something that progressives say to one another to justify restrictions on liberty. For—they say—it cannot possibly be the case that hate speech causes harm in this sort of way. Or if it does, the harm in question need not be taken seriously since it is not an outbreak of fighting or anything like that.

Occasionally one hears an acknowledgment that harm of this kind does exist, but the acknowledgment is coupled quickly with a refusal to be moved by it: hate speech is said to be protected constitutionally "not because we doubt the speech inflicts harm, but because we fear the

[34] *Id.* at 934–35.
[35] *Id.* at 933.
[36] *Id*.

censorship more."[37] This at least is candid, compared with the usual line of magical thinking: because speech is constitutionally protected, *it simply cannot be the case* that hate speech causes harm in this way.

The latter is an outcome-driven argument. And one hears a lot of it in these debates about free speech. It is said, for example, that degradation of the social environment of the sort that I have been talking about must not be counted as harm. To regard it as harm would be implicitly to undermine the status of free speech as a constitutional right. So, conversely, a full-blooded commitment to free speech as a constitutional right has to mean setting one's face against this sort of harm. I have found this a common response to my identification of a *dignitarian* dimension to the harm wrought by hate speech. It is said that the recognition of dignity as a legitimate interest is a sort of Trojan horse,[38] or (to vary the metaphor) it sets us on a slippery slope toward a European-style understanding of rights. Recognizing that dignity is at stake means admitting considerations of proportionality and balance into the rights equation. Such considerations, it is thought, undermine the distinctively principled character of American constitutional rights.[39] So, as good Americans, we must blind ourselves to the effect of hate speech on human dignity: that's the upshot of the first line of argument.

Or it is said that the model of causation I am using is inherently opposed to free speech as a constitutional right. No one who believes in free speech will be willing to accept that harm can accrue in the way that I have been suggesting. It is a mark of one's allegiance to the principle of free speech that one repudiates such a model of causation. This too is outcome-driven; we don't like the argument's outcome, so we infer that there must be something wrong with the structure of its causal analysis.

I think this sort of outcome-driven rejection of the slow-threat model of causation is unfortunate. Apart from anything else, discussions of it in the free speech arena almost invariably ignore the widely accepted use of such models elsewhere in public regulation. Professor Blasi's fine essay in this volume on the Holmesian response to claims about remote harm makes no mention whatsoever of the emissions or

[37] Michael W. McConnell, *You Can't Say That*, N.Y. TIMES (June 22, 2012), https://www.nytimes.com/2012/06/24/books/review/the-harm-in-hate-speech-by-jeremy-waldron.html.

[38] Guy Carmi, *Dignity—The Enemy from Within: A Theoretical and Comparative Analysis of Human Dignity as a Free Speech Justification*, 9 U. PA. J. CONST. L. 957, 958 (2007).

[39] *See* Neomi Rao, *On the Use and Abuse of Dignity in Constitutional Law*, 14 COLUM. J. EUR. L. 201 (2008).

other environmental examples.[40] It treats remote causation in the context of speech as though it were sui generis and as though the model of causation could be evaluated without reference to its use in any other domain or (to put it perhaps unfairly) as though these matters could be discussed without venturing beyond the confines of regulatory theory *circa* 1919.

And anyway, there is no reason for free speech advocates to be opposed in principle to this sort of causal model. On the contrary, something like it—the long-term and cumulative effect of lots of small-scale interactions—is often invoked in defense of free speech. Holmes's "free trade in ideas" conception—"the best test of truth is the power of the thought to get itself accepted in the competition of the market"[41]—relies on it. That conception invites us to consider speech acts not in and of themselves individually, but in a vast number of market-like interactions with hundreds of other speech acts, building upon and responding to one another, filtering out falsehoods and dead ends. Truth, if it emerges here, is a benefit but not a clear and present one; it is the result of complex sifting and accumulation.

Something similar is also involved on the negative side when thinkers like John Stuart Mill and his modern followers worry about the accumulation of conformist tendencies into "so great a mass of influences hostile to Individuality."[42] We hear an echo of it too in a recent open letter published in *Harper's Magazine* about the impact of "cancel culture":

> Editors are fired for running controversial pieces; books are withdrawn for alleged inauthenticity; journalists are barred from writing on certain topics; professors are investigated for quoting works of literature in class; a researcher is fired for circulating a peer-reviewed academic study; and the heads of organizations are ousted for what are sometimes just clumsy mistakes. Whatever the arguments around each particular incident, the result has been to steadily narrow the boundaries of what can be said without the threat of reprisal. . . .This stifling atmosphere will ultimately harm the most vital causes of our time.[43]

Each of the incidents alluded to might seem like an outrage to the freedom of those affected. But the letter is mostly concerned with the overall tendency of hundreds of such instances and the resultant

[40] *See* Vincent Blasi, *Holmes and the Marketplace of Ideas,* 2004 SUP. CT. REV. 1.

[41] Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

[42] JOHN STUART MILL, ON LIBERTY 132 (2nd ed. 1859).

[43] *A Letter on Justice and Open Debate*, HARPER'S MAGAZINE, July 7, 2020, https://harpers.org/a-letter-on-justice-and-open-debate/.

atmosphere. It ill behooves anyone who deploys the model of atmospheric accumulation in this way to denounce it as such when it is used by their opponents. I don't mean this as a simple *tu quoque*. One person might use the cumulation of small effects to make a case in favor of restrictions on speech, and one may also use the cumulation of *other* small effects to make a case against such restrictions. The fact that it is used on one side does not mean that exception cannot be taken to (e.g. the empirics of) its use on the other side. My point is simply that arguments of this kind should not be ruled out, out of hand.

All that said, it is important to acknowledge that these claims about atmospheric harms are not self-verifying. The power of the analogy with auto emissions that I began with doesn't make them true. Maybe the empirical evidence that hate speech leads to a slow accrual of harm is less convincing than it is in the auto emissions analogy. Jacob Mchangama, for example, says that "[w]hile we know for certain that the cumulative effect of automobile exhaust causes harmful pollution that can be averted with emission controls, we cannot with any degree of certainty know that the cumulative effect of hate speech results in 'environmental' harm avoidable by restricting free speech."[44]

It's a fair point. In other contexts, we might be quite skeptical of such claims about the accrual of harm. Suppose someone were to use the environmental analogy to justify prosecutions for sedition. (That, after all, was the real issue in the *Abrams* case.) Would defenders of restrictions on hate speech be willing to countenance defenses of the regulation of sedition—the prosecution of apparent trivialities like Mr. Abrams's "silly leaflet"[45]—on the ground that hundreds of such publications might add up over time to a genuine threat to national security? We must be sure that in our enthusiasm for considering this form of causation we do not give carte blanche to every alleged instance of the form.

Also, it is important to acknowledge that arguments for restricting hate speech based on slow causation are inherently vulnerable to the objection that *more speech* can always be inserted into the public debate to obviate long-term atmospheric threats. "If there be time to expose through discussion the falsehood and fallacies, to avert the evil by the processes of education, the remedy to be applied is more speech, not

[44] Jacob Mchangama, *The Harm in Hate Speech Laws*, HOOVER INST. POL'Y REV. (Dec. 2012), https://www.hoover.org/research/harm-hate-speech-laws (reviewing WALDRON, *supra* note 15).

[45] *Abrams*, 250 U.S. at 628.

enforced silence."[46] There is nothing in the mode of causation that I have been considering that rules out the possibility that countervailing speech might make such a difference. And there is nothing equivalent in the emissions analogy.

On the other hand, how we think about harm makes a difference in the way we treat the point about countervailing speech. The point about countervailing speech might easily be exaggerated, especially in a context where the hate speech in question is supposed to wreak its damage atmospherically rather than persuasively.[47] Defenders of free speech want the countervailing speech hypothesis to be true. But its common treatment as *just something effective to say* when concerns are voiced about the slow accumulation of harm hardly indicates that the prospect of that accumulation is being taken seriously.

[46] Whitney v. California, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring with Holmes, J.).

[47] *See also* Note, *Clear and Present Danger Re-Examined*, 51 COLUM. L. REV. 98, 106 (1951) ("[I]ndividuals vary in their degree of openmindedness. It would therefore seem that decisive significance can hardly be attached to the mere opportunity for free discussion, since the consequences of such an opportunity are variant and relatively unpredictable.").