Dangerous Speech

Jeffrey W. Howard

I. Introduction

On December 2, 2015, a husband and wife walked into an employee Christmas party in San Bernardino, California, and began shooting their semi-automatic rifles, killing 14 people. In his testimony to the Senate Judiciary Committee a week later, the FBI director assured senators that the couple were not members of any established terrorist organization. Instead they were inspired by "consuming poison on the internet"—specifically, through exposure to the normative arguments of the extremist cleric Anwar al-Awlaki, whose YouTube videos advocated the duty to kill Americans.[1]

On March 15, 2019, a man entered two mosques in Christchurch, New Zealand, killing 51 people. He was radicalized online, partly by a widely available manifesto written by the Norwegian white supremacist Anders Breivik, who himself had murdered 77 people in 2012. "We are against Islam," the manifesto announces, which "poses a mortal danger to the West," arguing that unless violent resistance is undertaken against Muslims—reviving the struggle begun in the Crusades—all non-Muslims will be enslaved. Encouraging violent attacks on civilians, the manifesto teaches that "it is better to kill too many than not enough, or you risk reducing the desired ideological impact of the strike." The manifesto clarifies: "It is not only our right but also our duty to contribute to preserve our identity, our culture and our national sovereignty by preventing the ongoing Islamisation."[2]

For all we know, had these individuals not encountered websites and online videos that spewed hate and encouraged terrorism, they would not have decided to become murderers. For all we know, had they not been incited by people peddling

[1] Al Baker and Marc Santora, "San Bernardino Attackers Discussed Jihad in Private Messages, F.B.I. Says," *New York Times*, December 16, 2015; Greg Miller, "Al-Qaeda figure seen as key inspiration for San Bernardino attackers," *Washington Post*, December 18, 2015.

[2] Anders Breivik, "2081 –A European Declaration of Independence," available at publicintelligence.net/anders-behring-breiviks-complete-manifesto-2083-a-european-declaration-of-independence". The Christchurch shooter's own manifesto has since been banned by the New Zealand government, sparking an outcry over the censorship of free speech. See "US says it will not join Christchurch Call against online terror," BBC News online, 15 May 2019, available at https://www.bbc.co.uk/news/technology-48288353

dangerous ideas, those 14 employees from San Bernardino and 51 worshippers in Christchurch would still be alive.

* * *

Is it morally permissible for the state to suppress expression to prevent the harms it risks inspiring? The U.S. Supreme Court insists that the answer to this question (or at least the answer to its constitutional variant[3]) is almost always *no*. Since the case of *Brandenburg v. Ohio* in 1969, the Court has affirmed sweeping protection for speech that advocates criminal conduct. Except for emergency cases in which the speech will cause imminent harm, it must be protected.[4] Because websites and online videos inciting murder typically do not cause harm imminently, their suppression would almost certainly be deemed an unconstitutional violation of the legal right to freedom of expression.[5]

This familiar American view stands in stark contrast to the position adopted in the United Kingdom, where encouraging terrorism is itself a crime, even when done implicitly through speech that "glorifies the commission or preparation" of terrorist acts. Further, it is a crime in the UK to incite harm through fomenting racial and religious hatred—for example, by expressing the view that citizens from certain groups are morally vile and so deserve to be attacked.[6] The British example is emulated in most liberal democracies' treatment of dangerous speech.

Who is right? While the British approach is globally dominant, the American view prevails in political philosophy. The idea that various forms of dangerous speech ought to receive capacious legal protection has been endorsed by, among many others,

---

[3] For some theories of constitutional interpretation, the questions of moral and constitutional permissibility are tightly connected; see Ronald Dworkin, *Freedom's Law: A Moral Reading of the American Constitution* (Cambridge, MA: Harvard University Press, 1996).

[4] *Brandenburg v. Ohio,* 395 U.S. at 447. This case involved a member of a Ku Klux Klan arrested for organizing and speaking in a televised rally in which he called for "revengeance" [*sic*] against black and Jewish Americans and their political supporters.

[5] This point is noted in Alexander Tsesis, "Terrorist Incitement on the Internet," *Fordham Law Review* 86, 2 (2017): 367-377, p. 369. One exception is for incendiary speech that simultaneously constitutes a "true threat"; see *Virginia v. Black* 538 U.S. 343 (2003). Such speech could permissibly be banned *qua* threat, but not *qua* incitement.

[6] See UK Racial and Religious Hatred Act 2006, available at http://www.legislation.gov.uk/ukpga/2006/1/contents, and UK Terrorism Act 2006, available at https://www.legislation.gov.uk/ukpga/2006/11/contents

Ronald Dworkin, Thomas Nagel, John Rawls, T.M. Scanlon, and Seana Shiffrin.[7] According to the prevailing view, the moral right to freedom of expression broadly includes incitement of wrongdoing within its protective ambit. And because the legal right to freedom of expression should be designed to track its moral counterpart, incitement merits legal protection. While there are many versions of this argument—corresponding to different accounts of what justifies the moral right to free speech—it constitutes the dominant philosophical strategy in defense of the extensive dangerous speech protections that the U.S. Supreme Court affirms in its *Brandenburg* decision.

My aim here is to debunk the philosophical orthodoxy on dangerous speech, and in so doing, defend a revised framework within which to conduct the debate. First, I show that the moral right to freedom of expression, properly interpreted, does not protect speech that incites clear violations of others' moral rights. Instead I argue that there is an enforceable moral duty to *refrain* from such incitement, a duty that shapes and limits the moral right to free speech itself. My thesis is that incendiary speakers can render themselves morally liable to coercion, and so, contrary to the orthodox view, are not necessarily wronged by such coercion. Notably, these enforceable moral duties do not map onto the strictures of the *Brandenburg* test, since, for example, they forbid incitement even when the incited harm will not occur imminently.[8]

Crucially, the mere fact that agents have an enforceable moral duty does not suffice to establish that the duty *should* be enforced, all-things-considered. In the second stage of the analysis, I concentrate on the issue of enforcement. As I shall show, many implausible claims in the free-speech literature, allegedly concerning the justification of the right to free speech, are rendered far more plausible once

---

[7] Dworkin, *Freedom's Law*, p. 200; Thomas Nagel, *Concealment and Exposure* (Oxford: Oxford University Press, 2002), p. 44; John Rawls, *Political Liberalism* (New York: Columbia University Press, 2005), pp. 336ff; T.M. Scanlon, "A Theory of Freedom of Expression," *Philosophy & Public Affairs* 1, 2 (1972): 204-226 (cf. "Freedom of Expression and Categories of Expression," *University of Pittsburgh Law Review* 40 (1978): 519- 550, which revises the earlier view but nevertheless indicates sympathy for *Brandenburg* on pp. 521, 539); Seana Shiffrin, *Speech Matters* (Princeton: Princeton University Press, 2014), p. 93, and "Speech, Death, and Double Effect," *New York University Law Review* 78 (2003): 1135-1185, p. 1139. Joshua Cohen also expresses qualified support for *Brandenburg* in "Freedom of Expression," *Philosophy & Public Affairs* 22, 3 (1993): 207-263, p. 236.

[8] A related issue concerns whether companies such as Facebook, Google, and Twitter are liable to be forced to remove dangerous content. The analysis here supports an affirmative answer, but I do not argue for that specific conclusion in any detail.

reconstructed as claims about enforcement—specifically, about why the enforcement of the duty to refrain from incitement would be all-things-considered unjustified.

To preview that argument, consider the prominent view that bans on dangerous speech are objectionable because these bans wrong *listeners*—insulting their autonomy, or depriving them of the valuable educative benefit of engaging with false views. While this view is traditionally understood as a justification for why a speaker's moral right to free speech protects incitement, I argue that it is more plausibly conceived as providing reasons not to enforce speakers' negative duties to refrain from incitement. Accordingly, such reasons must simply be weighed alongside the powerful reasons *to* enforce the duties in question, generated by the interests in protecting the prospective targets of incited harm. In this way, the real question at stake is whether the enforcement of a speaker's duty not to incite would be justified once factoring in all the effects of enforcing the duty on other people. I argue that other familiar arguments in the free-speech literature—e.g., concerning chilling effects, the counterproductive effects of driving dangerous speakers "underground", and the risks that statutes banning dangerous speech will be abused in the future—are also best understood as claims about the drawbacks of enforcement, to be factored into a morally weighted cost-benefit calculation. They are not considerations that justify the moral right to free speech itself.

How should this cost-benefit analysis proceed? The analysis is organized by a distinction between *narrow proportionality* in imposing harm, which is a matter of whether the target is morally liable to suffer that amount of harm, and *wide proportionality*, which factors in the effects of imposing that harm on non-liable people. The natural home for this distinction is in the ethics of defensive violence and warfare: even when an unjust aggressor is liable to be harmed in order to prevent her unjust attack, the wider collateral effects of the defensive harm on innocent bystanders may render the harm unjustified all-things-considered. I argue that this distinction has powerful payoffs when incorporated into free-speech theory. My contention is that while dangerous speakers are not typically wronged when subjected to legal restrictions on their incendiary expression, others may well be. If laws banning dangerous speech are likely to be politically abused, or counterproductively antagonize citizens and so increase danger, or cannot be written in a manner that is not objectionably over-inclusive, they may violate wide proportionality. Put in the terms of just war theory, the

collateral consequences of an otherwise just struggle against dangerous speech may simply be too great to justify it. Further, if enforcing the duty to refrain from incitement were reliably disproportionate in the wide sense, this could give legislatures reason to specify a *legal* right to free speech that is more protective than its purely moral counterpart. This would generate the result that citizens with no moral right to engage in dangerous speech would nevertheless enjoy a justified legal entitlement to do so. The possibility of an asymmetry between the justified law of a certain domain, and the underlying "deep morality" of that domain, is one of the fundamental findings of recent just war theory.[9] The argument here shows that such a phenomenon may arise in the free speech domain, as well.

Deploying insights from the ethics of defensive harm to refocus the debate on dangerous speech also helps us to identify a novel rationale for a common suggestion. While the dominant argument for protecting dangerous speech appeals to moral rights, a supplemental argument appeals to the preferability of *counter-speech* over coercion. I offer an interpretation of the counter-speech proposal that appeals to the necessity condition on the use of permissible coercion. If the danger of dangerous speech can be defused merely through *talking*, the use of coercion imposes harm excessively. I argue that while this is a plausible argument for why suppression of dangerous speech is sometimes morally wrong, it cannot vindicate the Supreme Court's requirement that incitement be permitted except when harm is imminent. There are many cases in which incited harms will not occur imminently, and yet in which the use of state coercion would satisfy the most plausible version of the necessity condition. Dangerous speech on the internet falls into this category.

My conclusions are simultaneously ambitious and modest. They are ambitious, because my aim is to propose a reorganization of the debate over dangerous speech, and in turn, freedom of speech itself. Once it is clear that the moral right to free speech does not entitle speakers to incite clear violations of others' rights, we can focus productively on the question of whether the duty not to incite should be enforced, all-things-considered, a matter that requires engagement with concerns of wide proportionality and necessity. But my aims are modest because I crucially leave open the answer to the

[9] The phrase "deep morality of war" traces to Jeff McMahan, "The Ethics of Killing in War," *Ethics* 114, 4 (2004): 693-733, p. 730.

question of what the right policy is. That is because concerns about wide proportionality are largely empirical, predicated on predictions about the effects of different laws and the tendencies of institutions to abuse their power. Still, even without the outcomes of that interdisciplinary work at hand, my arguments here suffice to show that, in a significant range of cases, speakers themselves are not wronged when the state enforces their duties to refrain from incitement. Provided concerns of wide proportionality and necessity could be satisfied, the state would enjoy far greater authority to suppress dangerous speech than the theoretical orthodoxy avows.

In Part II, I set out and specify the presumptive case for the duty not to incite the clear violation of others' rights. Part III assesses whether incitement might nevertheless be protected by the moral right to freedom of expression. Part IV re-evaluates the *Brandenburg* test in light of the preceding analysis. Part V focuses on considerations of wide proportionality, and Part VI turns to the matter of necessity.

II. Duty and Incitement

The contention I want to defend is that there is a moral duty to refrain from certain communicative speech that incites the clear violation of others' rights. The duty to refrain from such speech derives, I argue, from a general requirement not to endanger others without sufficient justification, and it is this underlying duty that fixes the contours of this category of wrongful speech. The claim that there is, in fact, a duty to refrain from incitement has received essentially no attention in the literature. Yet as I will show, its truth is of paramount relevance to the question of whether the moral right to freedom of expression protects dangerous speech.

The claim that there is such a duty could be resisted in one of two ways. The first approach grants that there is a general requirement not to endanger others without sufficient justification, but simply argues that our free speech interests can supply that justification in cases of incitement. On this view, the putative duty not to incite is vitiated, and incendiary speakers do no wrong. A second approach also grants the general duty of non-endangerment, and further concedes that we can derive from it a *bona fide* moral duty not to incite. Yet this second view holds that the interests underwriting free speech serve to render the duty *unenforceable*, granting the speaker an immunity against the enforcement of this duty—a so-called "right to do wrong". My

task in the next section (Part III) will be to show that neither approach succeeds; the duty not to incite is not vitiated by countervailing free speech considerations, nor rendered unenforceable by those considerations. In this section I set the stage for that analysis by sketching the presumptive case for the duty itself—its content and its contours.

What sort of dangerous speech is at issue here?[10] Specifically, the category of speech I seek to condemn is *speech that dangerously incites the incontrovertible violation of others' moral rights*. Clearly this requires some unpacking. First, simply as a conceptual matter, I take it that a speaker incites some conduct Φ if her speech (a) *advocates* her listeners to Φ *or* (b) *justifies* Φ. By *advocacy* of Φ, I refer to speech that involves encouraging, enjoining, imploring, beseeching, or persuading (though not forcing or threatening or inducing by offer of payment), be it *explicit advocacy* ("you should kill them") or *implicit advocacy* ("they are insects deserving extermination"[11]). By speech that *justifies* Φ, I refer to speech that supplies a putative (though not necessarily successful) justification for why Φ-ing is required, permissible, or otherwise to-be-done. Imagine a speaker who defends the permissibility of honor killings by arguing that rape victims have dishonored their families and for this reason lack the right not to be killed. Though he may not quite be *advocating* honor killings, this still qualifies as incitement. Note that in stipulating this definition, I am settling nothing as a normative matter.[12] I am simply declaring my topic.[13]

Second, I focus on violations of rights that are (to use a term of art) properly *incontrovertible*: it is beyond reasonable disagreement that they qualify as violations of a moral right. For example, while there may be reasonable disagreement over whether

[10] The term "dangerous speech" has been popularized in the excellent work of Susan Benesch and the Dangerous Speech Project; see https://dangerousspeech.org.
[11] The category thus includes the subset of so-called hate speech that takes the form of incitement.
[12] For further conceptual analysis of incitement, focusing on English law, see Joseph Jaconelli, "Incitement: A Study in Language Crime," *Criminal Law and Philosophy* 12, 2 (2018): 245-265.
[13] The framework I will defend can be marshalled to condemn other categories of dangerous speech. This includes reasonably uncontroversial cases, such as soliciting crime by offering payment, or other forms of crime-facilitating speech; for discussion, see Leslie Kendrick, "A Test for Criminally Instructional Speech," *Virginia Law Review* 91 (2005): 1973-2021, and Eugene Volokh, "Crime-Facilitating Speech," *Stanford Law Review* 57 (2005): 1095-1222. More controversially, I suspect that the dissemination of dangerous empirical falsehoods ("fake news") can also run afoul of the moral duty not to unduly endanger others and so may, in principle, be morally eligible for regulation; I am pursuing this line of argument elsewhere.

physician-assisted suicide or late-term abortion qualifies as a violation of the right to life, there should be no doubt that the vast preponderance of killings deemed murderous by the criminal law of most liberal democracies qualify as violations. The rationale for this restricted focus on incontrovertible violations—those beyond the scope of reasonable disagreement—is supplied by the moral right to free speech itself, and so will be brought into focus through the course of the next section.[14]

Finally, I refer not merely to speech that incites, but speech that (we have reason to think) *dangerously* incites. As I will explain, the wrongness of this category of speech traces to its dangerousness, i.e., to the way it endangers others through its foreseeable effects on susceptible listeners. As a result, incitement that endangers no one—what I will call *inert incitement*—simply does not qualify as wrongful (at least on grounds of dangerousness). The rationale for the focus on *dangerous* incitement, then, is supplied by the underlying duty that one violates by engaging in it.

* * *

I have sketched, and begun to justify, the contours of the category of dangerous incitement; the full justification depends on the arguments to come. To that end, we can now ask: why think that dangerous incitement, so specified, is wrongful at all?

Imagine a simple case involving three main parties—Adrian, Beatrice, and Cassandra. Adrian, addressing an audience including Beatrice, argues that members of Cassandra's religious group are vile scum who deserve to be killed. Beatrice subsequently attacks Cassandra. What has Adrian done wrong? The intuitive answer is that Adrian is failing to regulate his conduct by appropriate concern for the weighty, normatively significant interests of Cassandra (among others). Cassandra's weighty interest in life generates a moral duty for Adrian to refrain from conduct that frustrates that interest—a duty corresponding to a moral right held by Cassandra. While Cassandra clearly has a serious complaint against Beatrice for acting on Adrian's

[14] While I focus here on *serious* rights violations, such as murder, everything in the analysis to come applies to violations of any level of stringency. So, for example, I take it that there is a moral duty not to advocate pinching (under circumstances when this is dangerous), even though pinching is a minor (though indisputable) wrong. Whether the duty not to advocate pinching should be enforced is a matter for the proportionality analysis developed later on. My suspicion is that advocacy of trivial wrongs constitutes something like a *de minimis* offense under the law—i.e., offenses so minor that any legal enforcement is likely to be disproportionate (in either the narrow or wide sense or both).

exhortation to murder, it would be bizarre to think she did not also have a complaint against Adrian for his significant causal contribution to Beatrice's wrongful decision.[15]

Note that it would be no defense for Adrian to say to Cassandra's parents: "Beatrice is her own woman. She can make her own choices. I only *told* her to murder your daughter; it was up to her whether she actually *did it*." From their perspective, the fact that Adrian's foreseeably dangerous words operated through a responsible intervening agent, as opposed to some other force, is largely immaterial. Adrian cannot absolve himself of blame by pointing to the undeniable fact that the intervening agent Beatrice is a responsible adult, fully morally culpable for her decision to harm Cassandra. Even if Beatrice were a robot, and Adrian knew that uttering a certain word (e.g., "bingo") would risk triggering the robot to attack Cassandra (i.e., would increase the likelihood of the attack by the same probability as in the case where Beatrice is an agent), the grounds for objecting to Adrian's speech would be broadly the same.[16] Indeed, it may even be worse to cause wrongdoing through another moral agent.[17] Thus the mere fact that an intervening agent's conduct served as the bridge between Adrian's conduct and the harm that befalls Cassandra cannot morally launder his conduct, rendering it innocuous. They are both to blame.

Of course, when Adrian incites Beatrice to kill Cassandra, and Beatrice acts accordingly, the fact that Beatrice is the person who ultimately takes Cassandra's life in the final moment may suffice to establish that Beatrice, and only Beatrice, is a genuine murderer. But that is wholly compatible with the crucial contention that Adrian is *complicit* in the murder—he is an accomplice—for he culpably acted in a manner that causally contributed to the murder itself.[18] Just as the interests undergirding the right

---

[15] I think that *Beatrice* has a complaint against Adrian, too, for attempting to subvert her moral capacities; see my "Moral Subversion and Structural Entrapment," *The Journal of Political Philosophy* 24, 1 (2016): 24-46.

[16] See the insightful discussion in Larry Alexander, "Incitement and Freedom of Speech," in David Kretzmer and Francine Kershman Hazan (eds.), *Freedom of Speech and Incitement Against Democracy* (The Hague: Kluwer Law, 2000), pp. 114ff. Alexander argues—mistakenly, in my view—that incendiary speech is always fully protected when the intervening agent is a responsible agent (thus going even further than *Brandenburg*), but not when the intervening agent is non-responsible.

[17] Here I largely follow the view on intervening agency defended in Victor Tadros, "Permissibility in a World of Wrongdoing," *Philosophy & Public Affairs* 44, 2 (2016): 101-132.

[18] For the causal conception of complicity, see Chiara Lepora and Robert E. Goodin, *On Complicity and Compromise* (Oxford: Oxford University Press, 2013), and John Gardner, "Complicity and Causality," in *Offences and Defences*: *Selected Essays in the Philosophy of Criminal Law* (Oxford: Oxford University Press, 2007).

to life require us to refrain from murder, they also require is to refrain from *causally contributing* to murder—whether by inciting the murderer, or by knowingly driving the murderer to the murder scene, or by selling him weapons when it is obvious what he plans to do with them.

Note also that Cassandra's complaint against Adrian stands even if Beatrice resists Adrian's exhortations.[19] If Beatrice is successfully moved by Adrian, Cassandra's complaint may strengthen; that is an unavoidable implication of the phenomenon of moral luck. But even if Beatrice ends up doing nothing, Adrian wrongs Cassandra simply by engaging in conduct that *risks* causally contributing to the violation of her rights. The duty not to engage in dangerous speech, then, plausibly derives from a more fundamental duty, which I will call *the duty of non-endangerment.* This is a duty not to risk causally contributing to the incontrovertible violation of others' rights, unless one has a sufficient justification for doing so.

How could one have sufficient justification for engaging in conduct (such as speech) that risks causally contributing to the incontrovertible transgression of others' rights? Uncontroversially, if it is often wrong to impose risks of harm on others, it can be permissible to do so in certain cases. For example, it can be permissible to impose risks in cases in which those exposed to the risks have consented to them, or (perhaps) in which they benefit sufficiently from the risk-generating activity. Clearly those endangered by incitement typically do not consent, and (as I will argue) it is dubious that they benefit. But there is another kind of case in which risk-imposition can be morally justified: namely, cases in which the activity generating the risks is of sufficient moral significance to justify imposing such risks as an unintended, proportionate side-effect.

To see this possibility, consider the following case:

> *Atheist provocation*. Sonya advances sophisticated arguments for the truth of atheism, hoping to enlighten her audience. In an enraged response, a group of religious fanatics initiate a violent riot, murdering Sonya, her associates, and numerous other innocent people.

Now consider a distinct case:

---

[19] For further discussion on this point, see Jaconelli, "Incitement," pp. 248-250.

Atheist incitement. Tonya advances sophisticated arguments for the duty to murder atheists, inspiring her listeners to kill a group of them.

Part of the difference between these cases obviously involves the intention of the speakers; Tonya means to cause her listeners to kill, whereas Sonya does not.

But there is another crucial difference between the two cases. In Sonya's case, she is engaging in activity that she, intuitively, has every presumable moral right to engage in: the promulgation of her convictions on matters of religious truth. She has significant and weighty reasons to engage, and to be free to engage, in speech of this kind. In contrast, I am convinced that Tonya lacks such reasons to engage in her incendiary speech. An intuitively plausible response to Tonya's dangerous speech, if it's the only way to prevent the deaths of many innocent people, is to suppress it, for when it is suppressed, nothing of moral significance is lost. But the intuitively plausible response to Sonya's valuable speech is not suppression, but rather, *protection* for her and the other prospective victims of those listeners who are unreasonably provoked by it to kill. Provided such protection is available, it would clearly be wrong to respond to Sonya's speech by suppressing it.[20]

These claims simply assert what I take to be the intuitive difference between Sonya and Tonya. The idea that Sonya's pro-atheist speech merits significant protection is not controversial. What is controversial is whether Tonya's advocacy of murder also merits protection. I have just stipulated that Tonya's speech lacks value, such that there is no justification for engaging in it when it endangers others, but this is to declare precisely what I have the burden of showing. In the next section, accordingly, I turn directly to the question of whether speech that incites the clear violation of others' rights is nevertheless protected by the moral right to freedom of expression, as so many contend it is.

[20] What if police protection were not available? Were Sonya's life the only life endangered, I think she is permitted to provoke her own murder (though this would not amount, of course, to waiving her right to life). Given that others are endangered, too, matters are more complicated. On one view, the unavailability of police protection means that her legitimate speech may be permissibly silenced temporarily by the state in order to protect others—but this would *pro tanto* wrong her even if it were all-things-considered justified. For related discussion, see Timothy Garton Ash, "Defying the Assassin's Veto," *The New York Review of Books*, February 19, 2015, and *Free Speech* (London: Atlantic Books, 2016), pp. 130ff.

III. The Right to Incite

My task is to audit the arguments that purport to justify the moral right to freedom of expression, and to assess whether these arguments enjoin the protection of dangerous incitement within that right's protective ambit. In each case, I will show that they do not. The moral right to freedom of expression, properly understood, simply does not protect speech that incites the incontrovertible violation of others' rights. It neither justifies dangerous incitement, thereby vitiating the duty to refrain from incitement that I have defended; nor does it render such a duty unenforceable.

I will consider four central arguments for free speech.[21] The first (and longest) discussion focuses on the development and exercise of our moral powers. The second focuses on Seana Shiffrin's theory, which stresses the indispensable role of linguistic communication as a mechanism of access to others' minds. The third focuses on democracy, and the fourth discusses listener autonomy. Note that this discussion focuses on whether dangerous speech as I have specified it is protected by the moral right to freedom of expression; my discussion is agnostic about whether other forms of putatively harmful speech are protected.[22]

1. Free Speech and the Moral Powers

Consider, first, the idea that free speech is justified by the reasons we all have to value the development and exercise of our moral capacities. Freedom of speech matters, on this view, because we cannot adequately become reliable moral agents, reasoning about what morality requires and discharging its demands, in the absence of circumstances enabling free and open communication.[23] This argument traces to Rawls's specification of free speech as one of the *basic liberties*, those "essential social conditions for the

[21] Here I follow the standard methodology of free speech theory, which is to test whether a given category of expression is protected by working through the underlying free speech justifications. For another example of this method in a related context, see Caleb Yong, "Does Freedom of Speech Include Hate Speech?" *Res Publica* 17 (2011) 385-403. See also my "Free Speech and Hate Speech," pp. 96-100.

[22] I am thus exploring only the indirect relation between speech and harm, via audiences, rather than direct harm—in the form of harassment or other direct attacks on people's dignity of the sort that concern Jeremy Waldron in *The Harm in Hate Speech* (Cambridge, MA: Harvard University Press, 2012). For related discussion, see Frederick Schauer, "The Phenomenology of Speech and Harm," *Ethics* 103, 4 (1993): 635–653.

[23] Seana Shiffrin, *Speech Matters* (Princeton: Princeton University Press, 2014), pp. 91-92.

adequate development and full exercise of the two powers of moral personality over a complete life."[24]

Here I am focused in particular on the role of free speech for the development and exercise of the *first moral power*, or sense of justice.[25] I will largely set aside what Rawls calls the *second moral power*—the capacity to frame, revise, and pursue a conception of a good life—for the following reason. Conditions of open communication are uncontroversially central to the development and exercise of the second moral power—enabling agents to proselytize their views of the good and learn who else shares their views, and deliberate critically with others about matters of goodness and truth.[26] The crucial question is whether agents can appeal to the content of their conceptions of the good to justify inciting the violation of others' rights. Uncontroversially, the answer is no.[27] The sheer fact that one's religious doctrine, for example, enjoins the murder of infidels neither entitles one to murder infidels, nor to convince others to murder infidels. As Rawls influentially held, any moral rights tracing to citizens' second moral power must concern their pursuit of what he terms "permissible conceptions"—"comprehensive doctrines the pursuit of which is not excluded by the principles of political justice."[28] If I am correct that we are all presumptively under a duty to refrain from inciting clear violations of others' rights, as I have argued, the mere fact that one's comprehensive doctrine enjoins one to do so is immaterial.[29] That does not mean that there could not be *other* grounds for justifying the freedom to incite (either vitiating the presumptive duty not to incite, or rendering it robustly unenforceable); in what follows

[24] Rawls, *Political Liberalism*, p. 293.
[25] *Ibid.,* pp. 47ff.
[26] On these points, see Cohen, "Freedom of Expression," pp. 224, 228-229, and Joseph Raz, "Free Expression and Personal Identification," in *Ethics in the Public Domain: Essays in the Morality of Law and Politics* (Oxford: Oxford University Press, 1995).
[27] It is not *entirely* uncontroversial. Cf. C. Edwin Baker, "Scope of the First Amendment Freedom of Speech," *UCLA Law Review* 25 (1978): 964-1040, who defends the importance of self-expression even for those with deeply nefarious views. For criticism of Baker's argument in the context of hate speech, see Waldron, *The Harm in Hate Speech*, pp. 144ff. I take the Rawlsian argument summarized here to be an adequate objection to Baker; see my "Free Speech and Hate Speech," pp. 97-98.
[28] Rawls, *Political Liberalism,* p. 180.
[29] A similar insight arises in the debate over whether to grant exemptions to citizens whose comprehensive doctrines are burdened by (otherwise) just legislation. Cécile Laborde argues that exemption claims that are morally abhorrent (such as exemptions from the murder laws for those seeking to sacrifice infants) are denied consideration *ab initio*, since they have no weight; see *Liberalism's Religion* (Cambridge, MA: Harvard University Press, 2018), p. 209.

my aim is to explore precisely such grounds. But this argument suffices to show that merely pointing to the demands of one's comprehensive doctrine is insufficient.[30]

Speakers' interests

So let us focus on the argument for free speech that considers its role in facilitating the development and exercise of the sense of justice. Does freedom of speech, so justified, protect incitement?

Start with the idea of *exercising one's sense of justice*. This has both a reflective and a practical component. The reflective component concerns the task of reasoning about what one's duties of justice are. The practical component concerns the task of acting in accordance with the conclusions of one's reasoning. Suppose an agent does an awful job on both counts. We could say that she is exercising her sense of justice poorly, or simply that she is not exercising it at all. With respect to my topic here, I doubt much hangs on how we describe it. So we can simply opt for this pair of necessary and sufficient conditions: an agent exercises her sense of justice just in case she (a) reasons adequately about what her moral duties are and (b) acts accordingly. Accordingly, an agent fails to exercise her sense of justice whenever she reasons inadequately about what her duties of justice are, or fails to act accordingly.[31]

Note that to say that an agent exercises her sense of justice when she *reasons adequately* is not to say that she *reasons perfectly*, given the fact of reasonable disagreement; one reasons adequately just in case one reasons to a conclusion within the ambit of what is reasonable. This is the principal rationale for the aforementioned restriction on what I have termed the *incontrovertible* violation of others' rights. If there is reasonable disagreement about whether a certain act violates a right, there will accordingly be reasonable disagreement about whether there is a duty to refrain from

[30] Jonathan Quong rules out the protection of hate speech on precisely these grounds. Assessing whether an act is protected by a moral right, he argues, requires us "to ask whether the particular act that is alleged to be protected by a right is consistent with the overall moral ideal which the system of rights is meant to uphold."[30] See Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2010), p. 308. A similar idea is briefly expressed by Jeremy Waldron, "Rights in Conflict," *Ethics* 99, 3 (1989): 503-519, p. 518.

[31] When these come apart, we can disaggregate the claim, and refer to the respective exercise of the reflective and practical components.

encouraging that act.[32] Suppose there is reasonable disagreement about whether physician-assisted suicide violates the right to life. Speech that advocates physician-assisted suicide, then, does not qualify as inciting an incontrovertible rights violation, and so is not to be condemned by the argument here. Speakers' interest in exercising the sense of justice entitles them to express reasonable views about what justice requires.

But do we exercise our sense of justice by inciting the *incontrovertible* violation of others' rights? Consider our central example, speech that incites the incontrovertibly unjustified killing of an innocent person. It is very difficult to see how speech advocating or justifying such killing constitutes (a) adequate reasoning about one's moral duties. Similarly, it is hard to see how speech encouraging others to engaged in unjustified killing (b) instantiates successful moral reasoning in action. After all, as I argued in the previous section, we presumptively have a duty *not* to incite such killing. When we do so, our sense of justice has failed. Speakers do not exercise their sense of justice when they engage in such incitement.

So much for whether incitement constitutes an exercise of speakers' sense of justice. But what about the possibility that incitement facilitates the *development* of speakers' sense of justice? Consider the Aristotelian idea that the way to develop one's sense of justice, like any capacity, is to practice exercising it.[33] But if inciting murder does not qualify as a genuine exercise of the sense of justice, engaging in it cannot

---

[32] What constitutes a reasonable disagreement? Readers can help themselves to their preferred view. I opt for the notion of reasonable disagreement that is fairly mainstream in liberal political philosophy. On this view, a reasonable disagreement is one in which there are sufficiently plausible arguments on both sides. Following Rawls's influential discussion of the burdens of judgment, reasonable disagreement occurs when citizens who "share a common human reason" and hold "similar powers of thought and judgment"—i.e., who are similarly equipped to "draw inferences, weigh evidence, and balance competing considerations" and have access to the same high-quality evidence—nevertheless disagree (*Political Liberalism*, p. 55). In this same spirit, Christopher McMahon writes: "[T]he position taken by a party to a disagreement is reasonable if and only if it is or could be the product of competent reasoning." He continues: "Reasoning is competent when it is carried out in awareness of all the relevant considerations, the cognitive capacities exercised in extracting conclusions from the relevant considerations are appropriate, and these capacities are functioning properly" (*Reasonable Disagreement: A Theory of Political Morality* [Cambridge: Cambridge University Press, 2009], p. 8). In the context of reasonable disagreement about liberal justice, then, a disagreement counts as reasonable just in case each side of the disagreement counts as a comparably plausible interpretation of the requirements of liberal justice—understood as the ideal of persons as free and equal participants in fair social cooperation. For further discussion of this idea, see Quong, *Liberalism Without Perfection*, Chapter 7, as well as my "The Labors of Justice: Democracy, Respect, and Judicial Review," *Critical Review of International Social & Political Philosophy* 22, 2 (2019): 176-199, p. 187, from which this footnote draws.

[33] Aristotle, *Nicomachean Ethics*, Book II, Chapter 1.

facilitate the development of that capacity, for it does not constitute an instance of practicing it.

This argument is too quick. It seems at least possible that, in some cases, people develop their moral capacities through the experience of wrongdoing. "Doing wrong and later coming to regret it and learning from one's errors is not an unfamiliar path to moral self-constitution," Ori Herstein suggests. "Thus there is a benefit…to giving individuals the freedom to do wrong as a way of setting the conditions for subsequent contrition and as a step in self-directed moral development."[34] On this view, certain moral duties are unenforceable, because granting their bearers the choice to refuse to discharge them serves the development of their moral capacities. If incitement served the moral development of the speaker, then, her interest in such development could provide her with a Hohfeldian claim-right against the enforcement of the duty not to incite—a so-called "moral right to do wrong."[35]

The question we face, however, is whether the value of this development outweighs the disvalue of the harms inflicted on the victims of the relevant wrongdoing. Consider the claim that people should be allowed to *murder* others, since doing so serves their moral development. This is obviously implausible.[36] Even if we grant that there is *some* (instrumental) interest in murdering others when doing so serves one's moral development—say, because doing so triggers subsequent moral reflection—this interest is clearly not so great as to render unenforceable the duty not to murder. If murder is sufficiently egregious to outweigh the benefits of the wrongdoer's moral development, it would be surprising if *incitement* to murder were not, even if we grant that inciting murder is somewhat less objectionable than murder itself. [37] Certainly, it

---

[34] Herstein, "Defending the Right to Do Wrong," *Law and Philosophy* 31, 3 (2012): 343-365, p. 362.

[35] There is much debate on whether such an idea is substantively plausible or even coherent. I will assume *arguendo* that it is both plausible and coherent, as doing so puts me at a disadvantage by supplying an additional way in which free speech might protect incitement. For the initial defense, see Jeremy Waldron, "A Right to Do Wrong," *Ethics* 92, 1 (1981): 21-39, and another defense, David Enoch, "A Right to Violate One's Duty," *Law and Philosophy* 21 (2002): 355-384, For criticism, see William Galston, "On the Alleged Right to Do Wrong: A Response to Waldron," *Ethics* 93, 2 (1983): 320-324, Quong, *Liberalism Without Perfection,* pp. 307ff, and Renee Jorgensen Bollinger, "Revisiting the Right to Do Wrong," *Australasian Journal of Philosophy* 95 (2007): 43-57.

[36] Herstein acknowledges this, claiming that "one's autonomy interests in a right to do wrong are most likely never weighty enough to justify a right to highly egregious wrongdoing"—but never specifies what counts as egregious; see "Defending the Right to Do Wrong," p. 362.

[37] This seems to me an open question. On the one hand, encouraging murder simply creates an objectionable risk that someone will die, whereas engaging in murder actually kills, which is worse. On

would be surprising if the relevant threshold of egregiousness were located between them. Suppose Adrian is about to incite Beatrice to kill Cassandra, and Cassandra can suppress Adrian's speech by shooting him (and suppose this is the only available way to neutralize the threat). It would surely be perverse for Adrian to insist that Cassandra refrain from doing this, *in order to enable Adrian's moral development*. It is false, then, that speakers' interests in moral development could justify a right to incite.

Listeners' interests

There is a second set of development-oriented arguments, focused not on speakers' moral development, but instead on the interests of listeners. While the moral right to free speech most obviously concerns the right to *speak*, it is widely believed that the interests of listeners are relevant to the justification of the right.[38] One way to think about this idea, following Joseph Raz, is that while the moral right to speak one's mind is necessarily justified first and foremost by the interests of the speaker herself, the fact that listeners have interests in her having that right increases the stringency, or weight, of the right.[39]

There are potentially two ways in which listeners' interests in moral development could justify granting speakers the right to incite. First, exposure to dangerous views may *fortify* our moral capacities. On this view, we develop our moral capacities *by being tested*. Agents may have an interest in exposure to incitement, then, since such exposure may fortify our capacity to understand our moral duties and comply with them. In other words: by encountering those who enjoin us to do evil, we thereby become *less* likely to do evil.[40] But this is dubious. The purpose of encouraging a person to commit wrongdoing, after all, is to get her to commit wrongdoing. An inciter would have to be incompetent if the reliable result of his incitement was to *lessen* the likelihood that his

---

the other hand, if the probability of the listener's compliance is 100%, then inciting murder brings about the same result as murder but in a way that brings more culpable wrongdoing (the listener's) into the world, which may be worse. For the view that bringing more wrongdoing into the world is worse, see Tadros, "Permissibility in a World of Wrongdoing."

[38] E.g., Scanlon, "Freedom of Expression and Categories of Expression," pp. 524ff. For relevant discussion, see Leslie Kendrick, "Are Speech Rights for Speakers?", *Virginia Law Review* 103 (2017): 1767- 1809.

[39] Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), Chapter 10; Joseph Raz, "Free Expression and Personal Identification," *Oxford Journal of Legal Studies* 11, 3 (1991): 303-324, p. 150.

[40] For an argument offered in a similar spirit, see Lee Bollinger, *The Tolerant Society: Freedom of Speech and Extreme Speech in America* (Oxford: Oxford University Press, 1988).

listeners would engage in the incited course of action. Thus the claim that human beings, as such, have a standing interest in exposure to incitement, *because* it is fated to help them become motivationally resilient, is suspect.

There is a second version of this argument that is more plausible, which focuses not on the motivational benefits of exposure to sinister speech, but rather the epistemic benefits. On this familiar view, tracing to J.S. Mill, engagement with the most manifestly pernicious of views has value by enabling us to sharpen our understanding of the moral truth—"produced by its collision with error."[41] State suppression of speech that incites wrongdoing is objectionable because it deprives citizens of these valuable epistemic benefits, generated through exposure to an open "marketplace of ideas". Thus when the extremist cleric Abu Hamza encouraged his listeners "to bleed the enemies of Allah anywhere…to stab him here and there until he bleeds to death" since doing so is "the first stage of Jihad", he must be free to speak[42]—because his listeners have interests in exposure to such speech.

Even granting that listeners have *some* interest in exposure to the arguments for and against various forms of wrongdoing, we should doubt that this justifies a general moral right to engage in incitement. For starters, the interest in question is not plausibly an interest in exposure *simpliciter*, never mind *maximal* exposure. Rather, it is far more plausible to say that our interest is in some level of exposure sufficient to reap the epistemic benefits, in an environment conducive to the realization of those benefits—i.e., not an unregulated marketplace of ideas, but a structured environment that facilitates effective rational scrutiny of the dangerous speech in question.[43] The moral right to free speech, then, would protect incitement only under highly favorable conditions for effective criticism.

But there is a more fundamental problem with this argument, which sets the stage for my analysis later on. A moment ago I appealed to the claim, defended influentially by Raz, that the justification of a right to Φ must derive in the first instance

---

[41] Mill, *On Liberty*, Chapter 2, Section 1.

[42] Press Association, "The Preachings of Abu Hamza," *The Guardian*, February 7, 2006; at https://www.theguardian.com/uk/2006/feb/07/terrorism.world

[43] The idea that the truth is likely to prevail through a free, unregulated marketplace of ideas is by now one of the most criticized claims in free speech theory, such that few today seriously entertain it as a main justification of free speech. For a review of criticisms, see Paul H. Brietzke, "How and Why the Marketplace of Ideas Fails." *Valparaiso University Law Review* 31 (1997): 951-969.

from the interests of the putative right-bearer in Φ-ing (or in being free to Φ), though it can be bolstered by the interests of other parties.[44] One implication of this view is that if the putative right-bearer lacks an interest in Φ-ing (or in being free to Φ), then even if others have an interest in her having the moral right, this would not suffice to justify it. Now suppose, as I hope to establish, speakers simply have no serious interest in inciting the violation of others' rights; indeed, they are presumptively duty-bound to *refrain* from such activity. Thus even if listeners would value their exposure to incitement, that is not sufficient to justify the speaker's *right* to incite.[45]

This result is intuitive. Consider again Adrian, who is inciting Beatrice to kill Cassandra—say, because he thinks Cassandra's racial group deserves to be exterminated. Suppose a group of bystanders will reap epistemic benefits from hearing Adrian's racist argument (say, because it helps refresh their understanding of why they disagree). Suddenly, another bystander intervenes to stop Adrian from speaking. Surely, it would be perverse for Adrian to insist that he has the right to continue inciting Cassandra's murder, only to respond, when asked *why* he has such a right, by pointing to the educative benefits such incitement confers on the bystanders.[46]

That is not to say that the putative epistemic benefits of incitement are normatively irrelevant. If listeners truly have interests in exposure to incitement, then these interests would accordingly furnish genuine normative reasons. But these reasons do *not* serve to justify the speaker's moral right to incite. Instead, these reasons constitute considerations against the enforcement of the speaker's duty to refrain from incitement. Later, I will argue that these claims—in which the interests of third parties militate against the enforcement of speakers' duties—are best assessed through the criterion of *wide proportionality*, which concerns the effects of coercion on non-liable

---

[44] See also Herstein, "Defending the Right to Do Wrong," p. 348. For simplicity, this paper assumes the truth of interest-based theories of rights. Still, it seems that even those who reject interest-based theories of rights can appeal to the central idea under discussion here. After all, choice-based theories of rights appeal principally to the *choice-holder*'s claim to be free to make the relevant choice.

[45] This is consistent with the claim that for *other* speech—such as the promulgation of reasonable political and religious views—both speaker and listener interests contribute to the justification of the speakers' right to express such views. My point is that when it comes to a putative right to incite, listener interests alone are insufficient to justify such a right.

[46] There is a supplemental argument that supports this idea. For an agent to possess a moral right to *x*, the argument for her possession of that right must be "comprehensively justified" in G.A. Cohen's sense; it must be an argument that has moral force no matter who utters it. See Cohen, "Incentives, Inequality, and Community," *Tanner Lectures on Human Values* (1991), p. 279

parties. While the speaker has a duty not to incite, and so is not wronged by the enforcement of that duty (provided it is narrowly proportionate), third parties may have reasons to object to the enforcement—reasons which must be weighed against the reasons *to* enforce the duties in question. The Millean argument, I think, is better construed as a case of that sort. But once we see it as such, we should dispense with talk of the moral right to freedom of speech, and instead talk about whether it is proportionate to enforce duties to refrain from incitement. (I will return to this issue in detail in Part V.)

2. Free Speech and Mental Access

The most sophisticated recent argument for freedom of speech focuses on the interests that we all have as thinkers. On Seana Shiffrin's view, the central problem to which freedom of speech is a solution is the striking fact that "we lack direct access to the content of one another's minds."[47] Speech "provides the only precise mechanism by which one's mental contents may be conveyed to another mind, with all their subtlety and detail."[48] Without speech, our fundamental interests in having our thoughts known to others, and in knowing others' thoughts—and in achieving the complex epistemic and moral cooperation that such mutual knowledge makes possible—would remain significantly frustrated.[49] She thus calls for "open, unrestricted channels of communication" to enable the free expression of thought.[50] Accordingly, Shiffrin believes that freedom of speech protects a great deal of incitement.[51]

Shiffrin's theory is well-placed to defend dangerous incitement. Yet I want to argue for an interpretation of Shiffrin's theory that does *not* protect incitement. The argument begins by noticing that what is special on Shiffrin's view is not speech as such. It is *sincere* speech—veridical testimony, in which speakers are genuinely expressing what they think to others. Linguistic communication as a reliable mechanism for the

---

[47] Shiffrin, *Speech Matters*, p. 9.
[48] *Ibid.*, p. 10.
[49] *Ibid.,* pp. 88-94.
[50] *Ibid.,* p. 92.
[51] Shiffrin asserts that "incendiary speech" is protected in *Speech Matters*, p. 93. She also endorses the protection of incitement in her "Speech, Death, and Double Effect," *New York University Law Review* 78 (2003): 1135-1185, p. 113 (endorsing *Brandenburg* as "clearly correct")—though it remains unclear why she countenances *Brandenburg's* exception for imminently dangerous speech.

transference of thoughts is compromised when people misrepresent the contents of their minds. So, lying is wrong, on Shiffrin's view, because it hijacks a mechanism dedicated to the morally indispensable end of authentic communication, co-opting it for an improper purpose. In so doing, the liar "shows disrespect to our collective interest and duty to maintain reliable channels of communication".[52] Lying is so opposed to the goals that the right to free speech exists to protect, Shiffrin thinks, that it falls outside its ambit of protection.[53]

However, while Shiffrin believes there is a stringent prohibition on lying, she believes it can be permissible to lie, or misrepresent the contents of one's mind, in certain "justified suspended contexts."[54] These are cases, paradigmatically, in which one's interlocutor has no morally reasonable expectation that his questions will be answered truthfully. An interaction structured by norms of etiquette is one such context. *Pace* a conventional reading of Kant, Shiffrin also thinks that the interaction one has with a murderer on the hunt for his victim provides yet another justified suspended context, such that one is permitted to lie when asked by the murderer where his intended victim is hiding.

> Why is one permitted to lie to the murderer at the door? Shiffrin's answer is this: Given the compulsory ends morality supplies us, we could not reasonably use communication to further an evil end, and we cannot reasonably expect others to supply us with the reliable warrants necessary to do so.

Thus speech that constitutes "furthering an evil end", Shiffrin contends, cannot be reasonably expected, and should not be articulated, even when it is sincerely believed by the speaker. While she does not elaborate what speech "furthering an evil end" exactly involves, she has in mind speech that "would count as a form of assistance constituting accomplice liability."[55]

Earlier I argued that speech that incites the violation of rights is presumptively objectionable precisely because it is a form of wrongful complicity; it risks causally contributing to the wrongdoing of others. So if the value of veridical testimony is

---

[52] *Ibid.,* p. 24.
[53] *Ibid.*, pp. 116ff (though Shiffrin does not think lies should be banned all-things-considered).
[54] *Ibid.,* p. 16.
[55] *Ibid*., pp. 33-34.

vitiated, as Shiffrin thinks it is, in cases in which the testimony constitutes complicity with an evil purpose, the value of incitement seems null, as well. Even though the person supplying a murderer with the GPS coordinates of his victims is engaging in veridical testimony—even though she is genuinely sharing what is on her mind—his speech is unprotected.[56] So it goes, too, with the sincere inciter. There is, then, a reasonable interpretation of Shiffrin's powerful theory in which speech that incites the clear violation of others' rights is unprotected, despite its status as veridical testimony.

3. Free Speech and Democracy

A third influential argument for free speech appeals to the fundamental interests of citizens in a democratic society. While it would be implausible to suppose that this were the sole interest underwriting free speech, it is a significant one, justifying the special protection of "political speech", especially speech that criticizes state policy, that courts have provided.

This claim is ubiquitous in the American constitutional tradition.[57] "Shall we, then, as practitioners of freedom," asks Alexander Meiklejohn, "listen to ideas which, being opposed to our own, might destroy confidence in our form of government? Shall we give a hearing to those who hate and despise freedom, to those who would, if they had the power, would destroy our institutions?" The question is not rhetorical. "Certainly, yes!" he affirms. "It is the program of self-government."[58] Meiklejohn's suggestion is that it simply follows from what democracy *is*—a regime of self-government—that viewpoint-based restrictions on speech are ruled out as

[56] Shiffrin never explicitly states that conclusion, as the discussion of the murderer at the door occurs during her analysis of lying, not her analysis of free speech. But given that both discussions are part of a unified argument about the value of speech *qua* veridical testimony, this conclusion seems to me entailed by her view.

[57] For other views, see Cass R. Sunstein, *Democracy and the Problem of Free Speech* (New York: The Free Press, 1993), Robert C. Post, "Racist Speech, Democracy, and the First Amendment," *William and Mary Law Review* 32 (1990): 267–327, Steven J. Heyman, "Hate Speech, Public Discourse, and the First Amendment," in Ivan Hare and James Weinstein (eds.), *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2010), pp. 123–38, and James Weinstein, "Extreme Speech, Public Order, and Democracy: Lessons from the Masses," also in *Extreme Speech and Democracy*, pp. 23-61.

[58] Alexander Meiklejohn, *Free Speech and Its Relation to Self-Government* (New York: Harper and Brothers, 1948), pp. 65-66. See also his *Political Freedom* (New York: Harper and Brothers, 1960), pp. 26-27.

undemocratic. Echoing this sentiment further, Rawls contends: "to restrict or suppress free political speech…always implies at least a partial suspension of democracy."[59]

What, exactly, is the claim here? It will not suffice simply to appeal to what democracy *is,* as a conceptual matter.[60] The crucial question is whether we have reason to think that the normative ideal of democracy itself forbids speech restrictions. To answer that question, we must appeal to whatever interests are thought to justify democracy, and then explain whether these interests militate in favor of a conception of democracy in which speech restrictions are ruled out as objectionable. I will argue that once we pursue this strategy, we will see that the democratic case for protecting dangerous incitement is weak.

Consider three potential interests that are thought by different authors to justify democracy, evaluating whether they in fact enjoin the protection of dangerous speech. The first is what Niko Kolodny terms "substantive interests", namely, our interests in substantively just public policies that protect our rights and opportunities.[61] The instrumental justification of democracy depends on the claim that democracy better achieves just public policies, thereby satisfying citizens' substantive interests, than alternatives modes of governance.[62] Suppose, then, that this justification of democracy was the one we had in mind when specifying the free speech rights to which agents are entitled *qua* democratic citizens. If democracy is justified by our interest in securing just public policies, why would this interest militate in favor of protecting speech that advocated incontrovertibly unjust policies? If anything, it would be the other way around: our substantive interests would be better served if we lacked a right to dangerous speech.[63]

---

[59] Rawls, *Political Liberalism*, p. 354.

[60] This is the strategy in Eric Heinze, *Hate Speech and Democratic Citizenship* (Oxford: Oxford University Press, 2016), p. 5, who insists that, purely as a conceptual matter, any restriction on public debate "encroaches upon the elements that make the state a democracy." For criticism of Heinze, see Amanda Greene and Robert Simpson, "Tolerating Hate in the Name of Democracy," *Modern Law Review* 80, 40 (2017): 746–65.

[61] Niko Kolodny, "Rule Over None I: What Justifies Democracy?" *Philosophy & Public Affairs* 42, 3 (2014): 195-229, p. 200.

[62] See, for just one example, Richard Arneson, "Democracy is Not Intrinsically Just," In *Justice and Democracy: Essays for Brian Barry*, ed. Keith Dowding, Robert E. Goodin, Carole Pateman (Cambridge: Cambridge University Press, 2004), pp. 40-58.

[63] See the related discussion in my "Free Speech and Hate Speech," p. 99.

Most democratic theorists, of course, defend democracy non-instrumentally. So consider next the argument that democracy respects citizens' equal moral agency in the face of significant and reasonable disagreement about the proper trajectory of public affairs. This argument might appeal to the importance of respecting citizens' equal capacity to make moral judgments, or instead recognizing that citizens' interests have equal weight, or indeed to other fairness-based considerations.[64] Yet this argument, in its most plausible form, stresses the importance of respecting the views of all in the face of *reasonable* disagreement. There is no requirement to respect *unreasonable* views, which are beyond the pale.[65] As argued earlier, views that enjoin undeniable rights violations are not plausibly the product of adequate moral reasoning. While citizens' moral faculties are working satisfactorily in cases of reasonable disagreement—when there are good arguments on both sides of a debate—the same cannot be said of citizens who are defending incontrovertible injustice. Indeed, if the impetus for respecting citizens' moral agency flows from our recognition of them as agents of justice, endowed with the first moral power, then we rightly view they as duty-bound to *refrain* from activities that are morally incompatible with these duties. A moral prerogative to incite incontrovertible injustice is incompatible with the obligations of democratic citizenship.

A third notable argument for democracy holds that the purpose of democracy is to help constitute a society in which citizens relate to one another as social equals, rather than social superiors and inferiors.[66] This leads naturally to the worry that certain people who are disenfranchised, or lack the same rights to political participation as

---

[64] All three are invoked in the most influential argument developed along these lines, in Jeremy Waldron, *Law and Disagreement* (Oxford: Oxford University Press, 1999).

[65] Waldron, for example, explicitly frames his view as a response to "reasonable disagreements that are inevitable among people who take rights seriously"; see his "The Core of the Case Against Judicial Review," *Yale Law Journal*, 115 (2006): 1346-1406, p. 1369. I defend a variation of this argument in which the idea of reasonable disagreement is central in "The Labors of Justice," pp. 186-189. Another version of this idea—that democracy is a fitting response to reasonable disagreement—is defended in Thomas Christiano, *The Constitution of Equality* (Oxford: Oxford University Press, 2008), pp. 136ff, where he argues that judicial review is only acceptable when the "cores" of individual rights (understood plausibly as the aspects of the right beyond reasonable dispute) are threatened. Reasonable disagreement is also central to the democratic theory developed in Richard Bellamy, *Political Constitutionalism* (Cambridge: Cambridge University Press, 2007).

[66] See Niko Kolodny, "Rule Over None II: Social Equality and the Justification of Democracy," *Philosophy & Public Affairs* 42, 4 (2014): 287-336, and Daniel Viehoff, "Democratic Equality and Political Authority," *Philosophy & Public Affairs* 42, 4 (2014): 337-375.

others, will come to be regarded as socially inferior. Might restrictions on dangerous speech yield this result? Will a society that bans dangerous speech become an inegalitarian dystopia in which certain groups—the white supremacists and the jihadi extremists, to name but two—become marginalized?

This worry is misguided. For starters, even if restricting dangerous speech led to problematic relations of social hierarchy (between those who embrace liberal values and those who don't), this would not serve to vitiate or render unenforceable the duty not to incite incontrovertible rights violations. It would simply be another consideration to feed into the analysis (to be discussed shortly) of whether enforcement of the duty is all-things-considered justified.

More fundamentally, it is crucial to distinguish between the kind of problematic social hierarchy that arises when some section of the population is systematically condemned as inferior in status, and the crucially distinct phenomenon whereby citizens express disapproval of views they take to be evil. Darwall's familiar distinction between recognition and appraisal respect is paramount in this context; clearly we can *respect* white supremacists, in the sense of recognizing them as possessing an equal moral status and entitled to certain rights, without appraising their moral character or convictions in a positive light.[67] A common misreading of Rawls's view on so-called unreasonable citizens is that, just by dint of falling outside an overlapping consensus on liberal values, they would thereby forfeit certain rights.[68] But this interpretation is clearly mistaken.[69] Likewise, in a jurisdiction that banned dangerous speech, it's not true that certain citizens would have fewer rights than others. All citizens would have *the same free speech rights, with the same limits*. Consider an analogy with religious freedom. We don't think the right to religious freedom entitles people to engage in infant sacrifice—but that doesn't mean those whose religions enjoin them to sacrifice infants therefore lack the right to religious freedom. Everyone has the right to religious freedom, which is itself constrained by the same moral duties they owe to others. So it goes, too, with free speech.

---

[67] Stephen Darwall, "Two Kinds of Respect," *Ethics* 88, 1 (1977): 36-49.
[68] E.g., see Marilyn Friedman, "'John Rawls and the Political Coercion of Unreasonable People," in *Autonomy, Gender, Politics* (Oxford: Oxford University Press, 2003).
[69] Here I follow the views defended in Quong, *Liberalism without Perfection*, pp. 292ff.

There is a different version of this worry that appeals not to the value of democracy as such, but rather to the specific concern of democratic *legitimacy*. As Ronald Dworkin writes:

> it is illegitimate for governments to impose a collective or official decision on dissenting individuals, using the coercive powers of the state, unless that decision has been taken in a manner that respects each individual's status as a free and equal member of the community….The majority has no right to impose its will on someone who is forbidden to raise a voice in protest or argument or objection before the decision is taken.[70]

The suggestion, then, is that the legitimate authority of democratic decisions will be attenuated, at least with respect to those extremist speakers who find their speech suppressed. But as Jeremy Waldron has forcefully noted, the implication of this position is that the democratic legitimacy of most existing democracies to enforce the law—at least against citizens whose speech has been suppressed—is compromised. Yet it is deeply counter-intuitive to think that the British state may not use force to stop British terrorists' crimes due to the fact that terrorists' lacked a legal opportunity to communicate their support for such crimes in advance.[71]

Citizens who feel their speech has been objectionably suppressed may *think* the legitimate authority of the law over them has been compromised. But empirical beliefs about legitimacy, of the sort studied by political scientists, are of course different from the normative truth about whether a given institution does, *in fact*, command legitimate political authority. What grounds that? Clearly we cannot settle the basis of legitimate political authority here. So let's assume the popular thesis that a state's legitimate authority is grounded in the natural duty of justice.[72] Does the natural duty of justice militate in favor of a political arrangement in which citizens are free to incite incontrovertible rights violations through the political process—e.g., proposing legislation that perpetrates ethnic cleansing? Would a democracy's legitimate political

---

[70] Ronald Dworkin, "Forward," in Ivan Hare and James Weinstein (eds.), *Extreme Speech and Democracy* (Oxford: Oxford University Press, 2010), p. vii. A version of this legitimacy-focused position is endorsed by Corey Brettschneider, *When the State Speaks, What Should It Say?* (Princeton: Princeton University Press, 2012), pp. 75-78.

[71] Waldron, *The Harm in Hate Speech*, pp. 184-186.

[72] For various versions of this view, see Rawls, *A Theory of Justice*, 99-100, 115, 293-301, and 334; Anna Stilz, *Liberal Loyalty* (Princeton: Princeton University Press, 2009, Chapter 4; Quong, *Liberalism Without Perfection*, Chapter 4; Jeremy Waldron, "Special Ties and Natural Duties," *Philosophy & Public Affairs* 22, 1 (1993): 3-30.

authority, justified by appealing to this duty, be compromised if it denied citizens the opportunity to marshal support for incontrovertibly unjust legislation? It is difficult to see why it would. At any rate, an affirmative answer would need to recapitulate one of the arguments I have already canvassed and rejected.

So much for the democratic argument for protecting dangerous incitement. [73] Here's one final observation. It is often suggested that the mere fact that certain expression has political content, or is offered in public as part of democratic discourse, is *sufficient* to confer upon it significant protection.[74] Consider the claim, made by Kent Greenawalt, that the more public encouragement to crime is, the greater the free speech protection that applies, since the more it qualifies as public discourse. Yet if one has a duty to refrain from incitement, as I have argued here, a large audience (especially an impressionable one) may *amplify* the wrongness of such speech and so *strengthen* the case for prohibiting it, not weaken it. Greenawalt has it backwards.[75]

4. Free Speech and Listener Autonomy

Consider one final argument, according to which free speech is required by respect for listeners as autonomous. The most perspicuous statement of the view is offered by David Strauss: "the government may not suppress speech on the ground that the speech is likely to persuade people to do something that the government considers harmful. Put another way, harmful consequences resulting from the persuasive effects of speech may not be any part of the justification for restricting speech."[76] One version of this view

---

[73] Some might contend that respect for the democratic process requires that we refrain from encouraging the violation of legitimate laws (even ones we reasonably judge to be mistaken); we are allowed to advocate the repeal of such laws, but not their violation. But even if this argument were right, it would be a different argument from mine, as it would concern what communicative duties are derived from the more fundamental duty to respect the democratic process. My argument simply condemns speech that advocates incontrovertible rights violations. Whether we want to condemn *more* speech than that, on the basis of a different argument (such as this democratic one), is a further question).

[74] Cass Sunstein, *Democracy and the Problem of Free Speech* (New York: The Free Press, 1993).

[75] Greenawalt, *Speech, Crime, and the Uses of Language* (Oxford: Oxford University Press, 1989), p. 116. Likewise, I think Greenawalt is mistaken to suppose that, simply by dressing up one's incitement with the language of political conviction (so-called "ideological advocacy"), one can increase the protection offered (see his discussion on pp. 260-271). Insofar as the claims "The enduring maintenance of Enlightenment ideals requires that you kill Muslims" and "You should kill Muslims" are equally dangerous, they are equally wrongful. Indeed, the fact that the former is dressed up with political conviction may render it *more* dangerous, and so more wrongful.

[76] Strauss, "Persuasion, Autonomy, and Freedom of Expression," *Columbia Law Review* 91 (1991): 334-371, p. 335.

is cast as a deontological constraint according to which all state action must be compatible with a view of citizens as autonomous. On T.M. Scanlon's initial formulation of this view, "because a person must see himself as sovereign in deciding what to believe and in weighing competing reasons for action", it follows that "[t]he harm of coming to have false beliefs is not one that an autonomous man could allow the state to protect him against through restrictions on expression."[77] Another version hinges on the related but distinct notion that it is wrong to *express disrespect* for autonomy; as Ronald Dworkin puts the thought, the state "insults its citizens…when it decrees that they cannot be trusted to hear opinions that might persuade them to dangerous or offensive convictions."[78]

There is already a well-established objection to this position in the literature.[79] The central criticism of the view is that autonomous agents rightly recognize a tradeoff between exposure to dangerous messages and exposure to the harms those messages inspire.[80] Suppose citizens passed a law restricting dangerous incitement in order to protect themselves from the harms that speech could inspire. It would be puzzling for those citizens to feel insulted by their own efforts to protect themselves from harm.

Still, this argument is not decisive. Even if some citizens waive their claim to access such material—say, by supporting legislation that suppresses it—we could imagine other citizens (e.g., those who oppose the bans) insisting that their claim retains force. This is why we need a more fundamental reply to this argument than has been offered in the literature so far. This reply is precisely the argument I made earlier when discussing incitement's educative benefits for listeners. Even if it is true that banning certain expression sets back some interest held by listeners (in this case, the interest in respect, or in avoiding insult), this is not pertinent to the matter of whether the *speaker* has a moral claim to engage in the expression. As I have argued, speakers

---

[77] Scanlon, "A Theory of Freedom of Expression," pp. 215, 217.
[78] Dworkin, *Freedom's Law,* p. 200. See also Nagel, *Concealment and Exposure,* p. 44.
[79] Such that Scanlon has largely abandoned it; see "Freedom of Expression and Categories of Expression," p. 532.
[80] This point is made by Robert Amdur, "Scanlon on Freedom of Expression," *Philosophy & Public Affairs* 9, 3 (1980): 287-300, p. 299; a similar point is made with respect to hate speech in Susan Brison, "The Autonomy Defense of Free Speech," *Ethics* 108, 2 (1998): 312-339, p. 329. Perhaps this very criticism is what moved Strauss to offer a more moderate version of this argument than in Scanlon or Dworkin's initial formulation, contending—albeit without any elaboration—that the principle "can be overridden if the consequences of permitting the speech are sufficiently harmful"; see his "Persuasion, Autonomy, and Freedom of Expression," p. 360.

have no weighty interest in engaging in incitement. This fact undercuts any case for their possession of a moral right to incite.[81] So listeners' autonomy-based interests are best understood as claims to be fed into a wide proportionality calculation, concerning the effects of enforcing the duty not to incite on non-liable parties.

IV. Interlude: Reevaluating *Brandenburg*

I have argued that we have an enforceable moral duty to *refrain* from speech that incites the incontrovertible violation of others' rights, a duty which shapes our understanding of the moral right to free speech itself.[82] In the next two sections, I will turn to the matter of whether such a duty should be enforced, all-things-considered. But first, I want to return to the U.S. Supreme Court's standard, articulated in its *Brandenburg* ruling, of when it is constitutional to limit dangerous incitement. This test holds that so long as expression is not "directed to inciting or producing *imminent* lawless action" and is not "likely to produce such action," it must be protected.[83] Let's review the test in light of what we have learned.[84]

Note first that it is deeply puzzling that so many of the philosophers who defend *Brandenburg* appeal to the idea that the moral right to freedom of expression protects incitement. For if the moral right to freedom of expression genuinely protects incitement, why should it suddenly cease to do so as soon as the strictures of

---

[81] It might be suggested that because our fundamental rights, following Rawls, are specified under conditions of full compliance with morality, the fact that speakers have *some* interest in engaging in incitement is sufficient to justify a moral right to do so. After all, in conditions of full compliance, dangerous speech would *ex hypothesi* have no effect on listeners, who can be relied upon to do the right thing. Thus there are no countervailing reasons *in ideal theory* that militate against the existence of such a right. My reply is that even if speakers would have a moral right to incite fully trustworthy moral agents under those ideal conditions—just as they would have a moral right to sell guns to them, too—this fact does not mean that they have that right here and now in the real world. As T.M. Scanlon puts it: "Rights are not balanced, but are defined, or redefined, in the light of the balance of interests and of empirical facts about how these interests can best be protected" in "Rights and Interests," Kaushik Basu and Ravi Kanbur (eds.), *Arguments for a Better World: Essays in Honor of Amartya Sen: Volume I: Ethics, Welfare, and Measurement* (Oxford: Oxford University Press, 2008), p. 78. I thank an anonymous reviewer for raising this issue. It also helps me see that I was mistaken in earlier work to suggest that we must always first determine what our speech-involving rights are before then turning to establish what speech-involving duties we owe to others; I make this mistake in "Free Speech and Hate Speech," p. 100.

[82] That the moral right to free speech is less robust than commonly supposed does not diminish its status as special right, deserving a privileged status as one of our basic liberties; on this, I follow Leslie Kendrick, "Free Speech as a Special Right," *Philosophy & Public Affairs* 45, 2 (2017): 87-117.

[83] *Brandenburg v. Ohio,* 395 U.S. at 447. Emphasis added.

[84] For an instructive defense of *Brandenburg* at odds with much of what I say in this article, see Thomas Healy, "*Brandenburg* in a Time of Terror," *Notre Dame Law Review* 84, 2 (2009): 655-731.

Brandenburg are met? Consider again Shiffrin's view, according to which it is morally imperative that agents be free to communicate their sincere beliefs, however sinister. If this were true, why should the moral imperative in question lose force just so long as the *Brandenburg* conditions are met? Likewise, consider Scanlon's initial argument, predicated on respect for listener autonomy. If I am disrespected by bans on speech, then this disrespect surely obtains regardless of whether the speech satisfies the *Brandenburg* conditions or not.

Perhaps the idea is that, in emergency cases, in which violence is imminent, the right to free speech is permissibly infringed—*pro tanto* wronging the speaker (or listener) for the sake of preventing serious harm to others.[85] Yet if this were the view, the dangerous speaker who is silenced (and the listener who is thereby insulted) would be entitled to compensation and apology. This result is, needless to say, deeply counter-intuitive. Indeed, on the *Brandenburg* standard, the imminently dangerous speaker is eligible for *punishment*, which is difficult to square with the idea that he is owed compensation and apology.[86]

But the *Brandenburg* standard has deeper problems. The analysis I have offered helps us see that it does not reflect any coherent underlying normative principle concerning the duty to refrain from incitement. Start, first, with its requirement of intentionality. While *intending* to cause one's listeners to violate rights can aggravate blameworthiness (and, depending on one's view, wrongfulness), it is not necessary to violate the relevant duty. Even if one utters the phrase "Kill the Jews" *just for fun*, without *intending* to inspire anyone, it clearly remains wrongful if it nevertheless endangers Jews.[87] While we may hesitate as a general matter to *punish* people for merely

[85] Strauss, "Persuasion, Autonomy, and Freedom of Expression," p. 360.

[86] As one reviewer pointed out to me, one need not think that the speaker is *pro tanto* wronged in these cases; it may simply be that the reasons to prohibit the speech only become sufficiently strong in cases of imminent harm to outweigh the countervailing reasons to permit the speech. But then the question is why *only these cases* generate reasons of the requisite strength, given the dangerousness of *non*-imminent incitement.

[87] Baker suggests that insofar as expression aims merely to *express* one's views to others about what they should do, not to actually *influence* them, it merits protection; "Scope of the First Amendment," p. 994. A similar insight is defended by Greenawalt, *Speech, Crime, and the Uses of Language*, p. 123. This distinction seems to me to be fragile at best, and probably untenable.

reckless or negligent wrongs, intentionality is not necessary for violation of duty not to incite.[88]

Second, consider *likelihood*. The *Brandenburg* test is surely right to include some test for likelihood. Those whose speech poses no risk of harm at all may be offensive or disrespectful, but they do not violate the moral duty not to endanger the rights of others. Their speech needs to be *sufficiently* likely to inspire wrongdoing—but what is the standard of sufficiency? One candidate—the one suggested in *Brandenburg*—is *highly likely*. But this standard is implausibly demanding. In the run-up to the 1994 genocide in Rwanda, it is dubious whether any individual radio broadcast by the Hutu *interahamwe* militia led to any particular atrocity. The broadcasts aggregated to create a climate in which Hutus believed that they had a duty to engage in mass slaughter against Tutsis. Supposing the Rwandan government sought to protect the lives of Tutsi citizens, its hands strikingly would have been tied by the *Brandenburg* standard. It would not have been able to prosecute those behind individual radio addresses, whose dangerous effects materialized in a complex, aggregative fashion.[89] Or consider Abu Hamza, the North London cleric who encouraged his listeners to kill infidels. Even if the probability that someone in his cosmopolitan London audience would proceed to kill anyone was low, the gravity of the wrong incited—murder—makes the expression objectionable. The reason why is simple: it is wrongful to impose risks of egregious harms (such as death) on others for no good reason, even if the likelihood of the harm eventuating is small.[90]

---

[88] As Larry Alexander points out, the requirement of intentionality is more plausibly a condition placed by a normative theory of criminalization, rather than free speech; see "Incitement and Freedom of Speech," p. 108. See also Jaconelli, "Incitement," pp. 252-253.

[89] For a discussion of this point, applied to hate speech in particular, see Bhikhu Parekh, "Hate Speech: Is There a Case for Banning?", *Public Policy Research* 12 (2006): 213-223, p. 217. For a discussion of the complex mechanics by which speech can inspire violence, focusing on the Rwandan case, see Lynne Tirrell, "Genocidal Language Games," in Ishani Maitra and Mary Kate McGowan (eds.), *Speech and Harm: Controversies Over Free Speech* (Oxford: Oxford University Press, 2012).

[90] Consider the alternative standard articulated by Learned Hand, who wrote: "In each case [courts] must ask whether the *gravity* of the 'evil,' discounted by its improbability, justifies such invasion of free speech as is necessary to avoid the danger" (*Dennis v. United States,* 183 F.2d. 201, 212 (2nd Circuit, 1950). Without wishing to conclude anything now about what the right test is for courts to use (which was Hand's topic), Hand's test is a defensible way of thinking about the underlying moral duty itself. For arguments in defense of Hand's view focused on the appropriate test for courts, see Martin Redish, "Advocacy of Unlawful Conduct," *California Law Review* 70, 5 (1982): 1159-1200, pp. 1180-1182, and Alexander, "Incitement and Freedom of Speech," p. 113.

What about cases in which incitement poses no risk of inspiring harm at all?[91] Imagine I were to speak up at my family's Thanksgiving dinner, and suddenly encourage my family members to chop off the hands of the next left-handed person they meet. Knowing my family as well as I do, this would foreseeably endanger nobody; my family would be left puzzled and disturbed by the exhortation, but their probability of attacking left-handed people will not foreseeably increase. It is case of what I have called *inert incitement*. Is this still wrongful? One possibility is to say that while this speech-*token* does not endanger others, the speech-*type* of incitement *does* endanger others, and duties not to perform acts are indexed to types, not tokens. Thus it remains not merely disrespectful or offensive, but a *bona fide* violation of the duty not to endanger others through incitement. But it is not clear why this is so—why, in order words, our duties cannot admit of finer-grained differentiation between tokens of a common type, such that morality condemns certain tokens of the type but not others. If the wrong is the foreseeable endangerment of others, and the speech fails to endanger others, it cannot count as an instance of that specific wrong.[92] Here it is vital, as ever, to remember that context is crucially important to determining what is and is not dangerous.[93]

---

[91] Some doubt that dangerous speech actually has much of a causal role, at least within the contexts of liberal democracies. For the argument that we lack sufficient evidence that hate speech in particular is dangerous within developed democracies, see Heinze, *Hate Speech and Democratic Citizenship*, pp. 125ff. Crucially, one reason why Heinze seems to think that hateful speech isn't dangerous in developed democracies is that the police can adequately protect potential victims (p. 73). But this does not alter the dangerousness of the speech, any more than bulletproof vests render guns non-dangerous. Heinze does not dispute that speech can be dangerous outside of liberal democracies. For just one example of incitement's causal power, see David Yanagizawa-Drott, "Propaganda and Conflict: Evidence from the Rwandan Genocide," *The Quarterly Journal of Economics* 129, 4 (2014): 1947–1994. I thank David Skarbek for this reference.

[92] This does not necessarily mean that a law restricting dangerous speech would need to be crafted in such a way that only truly dangerous speech is proscribed. It is sometimes permissible to ban activity that, strictly speaking, is not wrongful, if it is unreasonable to expect the state to apply a more fine-grained legal standard. For example, Douglas Husak notes that ideally those speeding in their cars on roads they know are empty should be free to do so, since they don't foreseeably endanger anybody. But, he argues, it's simply too demanding to expect the state to distinguish between the morally innocent speeders and the morally culpable speeders, so it bans all of them; see his *Overcriminalization* (Oxford: Oxford University Press, 2008), p. 155) What's crucial to this argument is the assumption that the innocent speeders, with obvious exceptions, don't have morally significant reasons for speeding in the first place—and nor do inert inciters having morally significant reasons to incite. So, similarly, while inert inciters do not foreseeably endanger anybody, it may be too demanding for the state to craft statutes that reliably distinguish the citizen who inertly advocates murder from the citizen who dangerously advocates murder—so it just decides to ban all speech advocating murder (for example). If Husak is right, *this* kind of over-inclusiveness is not morally problematic (even if other kinds are, as I discuss in the next section).

[93] For discussion of the conditions that render speech dangerous, see Jonathan Leader Maynard and Susan Benesch, "Dangerous Speech and Dangerous Ideology: An Integrated Model for Monitoring and Prevention," *Genocide Studies and Prevention* 9, 3 (2016): 70-95.

Having now dealt with intentionality and likelihood, consider the third condition of *Brandenburg*: *imminence*.[94] My argument here suggests that there is a duty not to incite *even if* some period of time is going to elapse between the incitement and the inspired wrongdoing. Consider again Anwar al-Awlaki, the cleric whose YouTube videos exhorted viewers to kill infidels. When one reflects on whether one should post a video online encouraging viewers to engage in terrorist attacks, it does not seem to make a significant difference to the permissibility of the act how many minutes or hours or days will elapse between the posting of the video and the inspired terrorism. The duty holds regardless; a lack of imminence does not nullify it.[95] The idea that an incendiary speaker who incites his listeners to kill innocents can render his speech permissible simply by adding the caveat—"Don't kill them *now*; kill them *later*"—is implausible.[96]

Thus, on all three criteria, it is instructively clear that whatever the *Brandenburg* test is doing, it is neither specifying the conditions under which agents lose their moral right to incite, nor is it specifying the conditions under which agents have duties to refrain from incitement. The apparent lack of deep principle at work in the *Brandenburg* standard suggests it to be nothing more than an incoherent fudge.[97] But this would be too quick. The American view that we should largely permit dangerous speech secures greater plausibility when we turn away from the content of the duty not to incite, and instead consider the permissibility of *enforcing* that duty, all-things-considered. I turn to that issue now.

V. Speech, Harm, and Proportionality

---

[94] Notice that imminence seems to subsume likelihood, since if a crime is imminent, it is *a fortiori* likely. This point is also noted by Barendt, "Incitement to, and Glorification of, Terrorism," p. 458.

[95] Intriguingly, the U.S. law of criminal solicitation does not deem the temporal element important at all; even if one solicits a partner to commit a crime in a year's time, it can still constitute a criminal offence. Yet American criminal statutes banning criminal solicitation have never been officially challenged as inconsistent with *Brandenburg*. This observation is astutely noted by Alexander, "Incitement and Freedom of Speech," p. 114. My analysis clearly condemns solicitation as simply a version of incitement.

[96] I am thus sympathetic to practical proposals in the popular press to loosen the imminence requirement of *Brandenburg*; see, for example, Cass Sunstein, "Islamic State's Challenge to Free Speech," *Bloomberg Opinion*, November 23, 2015, available at https://www.bloomberg.com/opinion/articles/2015-11-23/islamic-state-s-challenge-to-free-speech; cf. Garton Ash, *Free Speech*, pp. 137-138.

[97] This is suggested in Alexander, "Incitement and Freedom of Speech," p. 118.

If agents have enforceable moral duties to refrain from incitement, this fact begins to justify state suppression of such speech. Such suppression serves to enforce the duties that these agents owe to others. There are two ways to suppress expression, corresponding to the two senses of the term *coercion*. The first operates simply by directly using force, or imposing harm, to prevent or stop it. For example, the state might gag a speaker about to engage in incitement, forcibly cupping his mouth, or shoot him in the leg to prevent him from carrying on with his dangerous address. Or it might jam the communication of an incendiary speaker on the radio. Or it might delete a hateful website (effectively destroying the speaker's property). The second operates by threatening criminal punishment, thereby providing speakers with a weighty prudential reason to refrain from undertaking the proscribed activity.[98] When the state criminalizes certain conduct, the point of doing so is precisely to authorize both forms of coercion, each a different technique of enforcing duty.[99]

Because speakers have enforceable duties to refrain from incitement, the state need not wrong them when it enforces these duties. In such cases, speakers are *liable* to coercion in some form.[100] Of course, speakers are not liable to suffer *any* amount of preventive coercion. Speakers are liable to suffer greater coercion in the service of preventing them from inciting grave wrongs, such as murder, than lesser wrongs, such as destruction of property. There is also the important question of whether the coercion that inciters are liable to suffer is the same amount of coercion that those they successfully incite are liable to suffer. But how to calibrate exactly how much, and what

---

[98] This distinction is flagged in Alexander, "Incitement and Freedom of Speech," p. 112.

[99] The justifications of defensive harm and punishment are distinct in various respects. But on any plausible view, there will be deep connections between these justifications. After all, when we standardly speak of the state's authority to use coercion, we are referring *both* to its authority to use direct force to prevent (harmful) wrongs, *and* its authority to threaten and impose punishment for the commission of these wrongs. For deontological deterrence theorists, the connection will be very deep, as in Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law* (Oxford: Oxford University Press, 2011), and *Wrongs and Crimes* (Oxford: Oxford University Press, 2016). But even retributivists, who view punishment as a special normative domain governed by considerations of moral desert, should recognize a deep connection, since it is culpable (harmful) wrongdoing that triggers *both* the permissibility of defensive force (in standard cases involving culpable aggressors) and the permissibility of punishment. The general point is simply this: the conduct that it is permissible to use force to prevent tends also to be the kind of conduct that it is permissible to punish, and the state's coercive power involves both phenomena.

[100] The idea that "a person is liable to be harmed for some goal if she has an enforceable moral duty to bear that harm for that goal" is defended in Victor Tadros, "Duty and Liability," *Utilitas* 24 (2012): 263. How much harm one is liable to bear is a further question.

deeper principle explains it, is largely the task of a full theory of proportionality—something I leave open for present purposes.[101] What matters is that inciters are liable to suffer *some* proportionate amount of coercion, and so need not be wronged by it.

Even if the wrongdoer is not wronged by the coercion, because she is liable, others, who are not liable, could be. I have already flagged the distinction between *narrow* proportionality, which is a matter of whether a wrongdoer is liable to suffer coercion in the service of preventing the wrong, and *wide* proportionality, which concerns the effects of such coercion on third-parties.[102] Suppose Beatrice is about to shoot Cassandra, and Cassandra's sole means of protection is to toss a grenade at Beatrice. Beatrice would not be wronged by the grenade blast, which kills her; it is narrowly proportionate. But if the grenade blast is likely to kill a hundred innocent bystanders, who have not forfeited their rights, it is nevertheless disproportionate in the *wide* sense, and thus all-things-considered impermissible.

I have already intimated that various concerns in the free speech literature are best understood in terms of wide proportionality. Consider the putative educative benefits of exposure to incitement for listeners. Given that speakers have no right to engage in such speech—indeed, they are duty-bound to refrain from it—I said that this consideration is far more plausibly understood as a reason to refrain from enforcing the duty in question. In this way, the epistemic benefits of dangerous speech for listeners simply figure as part of a wide proportionality calculation. This puts them in their proper place. For even if Mill is right that such epistemic benefits exist, we need to weigh these putative epistemic benefits (of exposure to anti-Semitic screeds and racist rants and pro-terrorist sermons) against the risks to life and limb generated by permitting them. Faced between the choice of protecting vulnerable people from murder, and offering others a valuable intellectual exercise, it does not seem to me a close call what

[101] Prevailing views index proportionality to the degree of moral responsibility of the wrongdoer. For a criticism of these views, alongside an alternative that indexes proportionality to the stringency of the right threatened, see Jonathan Quong, "Proportionality, Liability, and Defensive Harm," *Philosophy & Public Affairs* 43, 2 (2015): 144-173.

[102] See Jeff McMahan, *Killing in War* (Oxford: Oxford University Press, 2009), pp. 20-21, and "Proportionate Defense," in Jens David Ohlin, Larry May, and Claire Finkelstein (eds.), *Weighing Lives in War* (Oxford: Oxford University Press, 2017), pp. 135ff. See also Quong, "Proportionality, Liability, and Defensive Harm," p. 145n.

matters more. But one need not accept that specific point to see the general one: that this is a matter of wide proportionality.

My proposal is that a variety of familiar claims recurrently made in the free speech literature are also far better understood as claims about why it would be disproportionate in the wide sense to enforce the duty to refrain from incitement. I will briefly discuss three. First, consider the recurrent idea that the right to free speech should protect dangerous speech since dangerous speakers would otherwise be "driven underground"—i.e., they would still preach dangerous ideas, but covertly.[103] One problem generated by this effect is that it makes it more difficult to identify who the dangerous people in our midst are.[104] Insofar as we want to be able not simply to influence susceptible audiences, but try to persuade dangerous speakers to change their views, this aim would thereby be frustrated. Another problem generated by driving dangerous speakers underground is that it may make them *more* dangerous. By depriving them of a valuably cathartic legal outlet through which to vent their anger in a community of likeminded associates, it is possible that dangerous speakers will become *more* antagonistic toward the rights of others, and so possibly more likely to engage in violence themselves (even if their power to influence others has been largely neutered).[105] It may be, then, that *more* innocent lives are endangered by *banning* dangerous speech than by permitting it.[106] Whether this is true is an empirical question. My point is that it is not a question about the moral right to free speech itself. It is, rather, a matter of whether it would be wide-disproportionate to enforce the duty to refrain from dangerous speech.

Second, consider the idea that we shouldn't ban dangerous speech because doing so would *inevitably suppress legitimate speech*. There are two versions of this worry. First, it may simply be too difficult to author statutes restricting dangerous speech in a manner that isn't objectionably over-inclusive, coercing non-liable people. For example,

---

[103] For a recent defense of this claim, see Nadine Strossen, *Hate: Why We Should Resist it with Free Speech, not Censorship* (Oxford: Oxford University Press, 2018), pp. 143ff.

[104] This is a familiar point, made in Barendt, "Incitement to, and Glorification of, Terrorism," p. 453.

[105] Nancy Rosenblum, *Membership and Morals: The Personal Uses of Pluralism in America* (Princeton: Princeton University Press, 1998), pp. 254ff.

[106] A related worry is that bans against dangerous speech will simply make no difference, since the elusive all but guarantees their ineffectiveness. I am skeptical about this claim, but in any case, this is a concern about the success condition for the use of permissible force—an important issue I do not discuss here—rather than a worry about proportionality or necessity.

in our current far-from-fully-just societies, a statute banning all criminal advocacy, say, would restrict plenty of morally permissible speech. (Consider a speaker who encourages a gay couple to act on their love in a society that unjustly banned gay sex.) But second, even if we only proscribe genuinely wrongful speech—speech that agents are duty-bound to refrain from—such an effort may have an unfortunate side-effect on legitimate speakers. Citizens may self-censor in cases when they shouldn't have to—for example, out of fear that their legitimate criticism of a certain religion, say, will be (mistakenly) condemned as dangerous speech, or that their legitimate criticism of the government will be (mistakenly) condemned as support for terrorists who invoke the exact same criticism to rationalize their wrongful crimes. Whether speech laws genuinely have this so-called "chilling effect" is, as before, an empirical question, but it is a question to be examined within an assessment of wide proportionality.[107]

Third, consider another familiar claim in this area: we simply should not trust the government to regulate speech. Politicians are often unscrupulous, and human history is littered with examples of powerful leaders abusing speech restrictions to harm their political adversaries, suppressing legitimate speech.[108] Sometimes this is the result of moral ignorance; other times, malice. These concerns, grounded in the interests of those who would be unjustly coerced by misguided speech restrictions, generate reasons against the enforcement of the moral duty not to incite, for the following reason: the very instrument that enables the state to target liable speakers would invariably be deployed against non-liable speakers. On this view, a criminal statute suppressing dangerous speech is simply too dangerous an instrument to permit the state to have. These concerns, too, properly enter the analysis not at the level of determining what the moral right to free speech is, but rather in determining whether the duty to refrain from incitement should or should not be enforced, all-things-considered.

---

[107] For related discussion, see Frederick Schauer, "Fear, Risk, and the First Amendment: Unraveling the Chilling Effect," *Boston University Law Review* 58 (1978): 685-732; Strossen, *Hate*, pp. 99ff; and Leslie Kendrick, "Speech, Intent, and the Chilling Effect," *William & Mary Law Review* 54 (2013): 1633-1691.

[108] For arguments in this spirit, see Geoffrey Stone, *Perilous Times: Free Speech in Wartime* (New York: W.W. Norton & Company, 2004), Richard Epstein, "Property, Speech, and the Politics of Distrust," *The University of Chicago Law Review* 59, 1 (1992): 41-89; Vincent Blasi, "The Checking Value in First Amendment Theory," *American Bar Foundation Research Journal* 3 (1977): 521-649; and Strossen, *Hate*, p. 82.

Notice that result of that wide proportionality analysis need not be all-or-nothing—i.e., it need not conclude with a complete refusal to enforce the relevant duties. We might conclude that, for reasons of caution, we should adopt under-inclusive statutes (i.e., statutes that deliberately ban less speech than is, in principle, eligible for suppression). Or we might adopt a statute that only banned offences whose wrongness we had a high degree of confidence, such as murder, or that refrained from banning certain forms of violence against the government that might, in emergency conditions, be justified.[109] Or we might eschew the use of the criminal law entirely, deploying civil mechanisms instead to deter and compensate the harms caused by dangerous speech.[110]

The popular temptation to think that these concerns—counter-productivity, risks of abuse, chilling effects—are relevant to the moral right to free speech traces, I suspect, to the fact that they may be relevant to the proper specification of the *legal* right to free speech.[111] Because it is widely assumed that the legal right to free speech simply tracks the underlying moral right, it is natural to assume that what is relevant for one is relevant for the other.[112] But this need not be the case. If enforcing the duty to refrain from incitement is reliably disproportionate in a wide sense, this gives legislatures decisive reason to refrain from enacting statutes that enforce this duty. This would generate the result that citizens with no moral right to engage in certain speech would nevertheless enjoy a justified legal entitlement to do so.[113] As I mentioned at the outset, the possibility of an asymmetry between the justified law of a certain domain, and the underlying "deep morality" of that domain, is one of the fundamental findings of recent just war theory. My argument here has demonstrated that such a phenomenon arises with respect to free speech, as well.

---

[109] Consider Mill's (qualified) suggestion in *On Liberty*, Chapter 2, that "tyrannicide"—assassination of the king—must be broadly protected by freedom of speech. Similarly, insofar as violent revolution against the state may in certain extremely rare cases be justified, we may want to insist on a prophylactic measure whereby we protect all advocacy of violent revolution—or even all seditious speech—even though such a rule is clearly over-inclusive, protecting lots of liable speakers (e.g., unreasonable terrorists) who would not be wronged by restrictions on their speech.

[110] On this point, see Frederick Schauer, "Uncoupling Free Speech," *Columbia Law Review* 92, 6 (1992): 1321-1357.

[111] For example, this seems to be the underlying view in Strossen, *Hate*, pp. 37ff.

[112] Cf. Shiffrin, "Speech, Death, and Double Effect," p. 1184.

[113] This need not strike us as puzzling. Suppose that, in accordance with Rawls's difference principle, which permits inequalities that benefit the worst-off, the best tax system gave "self-seeking high-fliers" (in G.A. Cohen's memorable terms) a hefty salary. These individuals have no pre-institutional moral right to that salary, merely a justified legal entitlement. I am grateful to Tom Parr for this example.

VI. Counter-Speech, Necessity, and Imminence

Wide proportionality serves as one constraint on the enforcement of a moral duty; *necessity* serves as another. The condition holds, roughly, that no more force, or harm, than necessary to defuse a threat is permissible. So, if I can defuse the unjustified lethal threat you pose by shooting you in the leg, then I ought to do that rather than shoot you in the head, *ceteris paribus*. While the necessity condition arises most familiarly in discussions of defensive harm, it is plausibly understood as a general condition on the use of coercion—both in the sense of directly deploying force, and in the sense of threatening and imposing criminal punishment.[114] There is a lively debate on what the justification for this principle is and how it fits with the ideas of liability and proportionality, but we need not get entangled in that debate here.[115] Instead I merely want to argue for one powerful but ignored implication for this principle (whatever its underlying normative structure might be) for the debate on dangerous speech.

Consider a simple example. Adrian is again attempting to persuade Beatrice that she has an obligation to kill Cassandra. Dennis, a bystander, is the only person who can take action to stop this. Suppose Dennis could stop the murder through *counter-speech*—either by talking Adrian out of his nefarious plan to incite Beatrice, or by allowing Adrian to speak but then supplying Beatrice with convincing counter-arguments. The counter-speech strategy, we can stipulate, will have a 100% chance of

[114] Given that criminalization of wrongful conduct is the central mechanism by which the state enforces moral duties to refrain from that conduct—both by authorizing direct force to stop wrongdoers, and by threatening criminal punishment—this is unsurprising. The necessity condition on the permissibility of criminalization holds, roughly, that unless criminalization is necessary to achieve its purpose, we should opt for non-criminal modes of response. This position is sometimes encapsulated by the claim that the criminal law is "a last resort." How, exactly, the necessity principle for criminalization relates to the necessity principle for defensive harm—whether they trace to an overarching necessity principle for all coercion, as I suspect they do, or whether they are subtly distinct principles that are morally related in some different way—is a fascinating and under-theorized question that I cannot resolve here. For discussion, see Douglas Husak, "The Criminal Law as a Last Resort," *Oxford Journal of Legal Studies* 24, 2 (2004): 207-235, and "Applying *Ultima Ratio*: A Skeptical Assessment," *Ohio State Journal of Criminal Law* 2 (2005): 535-545.

[115] A crucial aspect of this debate concerns whether necessity is *internal* to liability, such that aggressors are only liable to whatever coercion (beneath the proportionality ceiling) is necessary to defuse their unjustified threat, or whether necessity is instead *external* to liability. For defenders of the former view, see McMahan, *Killing in War*, p. 9, and Seth Lazar, "Necessity in Self-Defense and War," *Philosophy & Public Affairs* 40, 1 (2012): 3-44. For defenders of (very different versions of) the externalist view, see Joanna Mary Firth and Jonathan Quong, "Necessity, Moral Liability, and Defensive Harm," *Law and Philosophy* 31, 6 (2012): 673-701, and Helen Frowe, "The Role of Necessity in Liability to Defensive Harm," in Christian Coons and Michael Weber (eds.), *The Ethics of Self-Defense* (Oxford: Oxford University Press, 2016).

saving Cassandra's life, at a low cost to Dennis. Suppose the only alternative for Dennis is to *kill* Adrian before he incites Beatrice; this will also have a 100% chance of success in rescuing Cassandra, at a similarly low cost to Dennis. Dennis kills Adrian. This seems to violate the necessity condition; Dennis could have stopped Adrian through *talking*, but instead needlessly deployed lethal force. In this case, even if lethal force would (otherwise[116]) be proportionate, it is gratuitously harmful.

This insight—that the use of coercion is sometimes objectionable, because it is unnecessary—supplies a novel justification for the familiar thesis that we should use *counter-speech*, rather than coercion, to respond to dangerous expression.[117] This idea is a familiar one, but the exact argument for this idea has never been entirely clear. For many scholars, the importance of counter-speech traces to the fact that coercively suppressing speech violates the moral right to freedom of expression. For them, counter-speech is the only permissible remedy. As I showed earlier, this argument fails, since restrictions on dangerous incitement do not violate the moral right to free speech. So how else might the Supreme Court's emphasis on counter-speech be vindicated? The possibility on which I am focused here is the following: even if the use of coercion to suppress dangerous speech met proportionality concerns, it might still violate the necessity condition.

This argument could, in turn, furnish a principled moral rationale for *Brandenburg* for which we have long been searching but not yet found. It would do so by justifying the Court's otherwise puzzling view that coercion is impermissible except in cases of imminent danger. The idea would be that only in cases of imminent danger is coercion necessary; in cases of non-imminent danger, it is unnecessary, since there is less harmful mechanism—counter-speech—for defusing the speech's danger. As far as I know, no one has ever suggested the necessity condition as the rationale for

[116] For certain internalists about necessity, unnecessary harm is always disproportionate; see Lazar, "Necessity in Self-Defense and War," pp. 17ff. This is why he deems proportionality and necessity as merely "superficially distinct"; see Lazar, "Just War Theory: Revisionists vs. Traditionalists," *Annual Review of Political Science* 20 (2017): 37-54, p. 44.

[117] The historical genesis of this proposal is Justice Louis Brandeis's concurring opinion in *Whitney v. California*, 274 U.S. 357 (1927), cited in *Brandenburg*. For recent defenses of counter-speech, focusing on the role of the state, see Brettschneider, *When the State Speaks*, and Strossen, *Hate*, pp. 158ff. See also my "Terror, Hate, and the Demands of Counter-Speech" *British Journal of Political Science* (forthcoming).

Brandenburg's imminence requirement. It is, I think, the most promising rationale available.

Still, here are three reasons to doubt that a suitably specified necessity condition would vindicate the imminence requirement. First, it is widely believed that while imminence and necessity often go together, then often can come apart. In both the ethical and legal scholarship on defensive harm, it is widely recognized that imminence is merely a proxy for necessity. As David Rodin puts it, an imminence requirement "is simply the application of the necessity requirement subject to epistemic limitations" since "we cannot know with the required degree of certainty that a defensive act is necessary until the infliction of harm is imminent."[118] Yet it is not plausible that the necessity condition always requires that we wait until a threat is imminent. [119] Consider an example from Helen Frowe:

> *Isolation*. Murderer…wants to kill Victim. He has chased Victim into the desert, where Victim is hiding in an abandoned, but well-fortified, building. There are no means of communication with the outside world. Murderer is waiting outside, shouting to Victim that he doesn't plan on going anywhere. Victim knows that it will take Murderer several days to penetrate the building's defences, but that once he does, Victim's chances of a clear shot at him will be much smaller. [120]

The intuitive response, according to Frowe, is that Victim is not required to wait for imminence at the cost of significantly lessening her chances at successfully defending herself. Or consider an example in the dangerous speech context:

> *Incommunicado hitman*. Susan is about to send an email to her hitman, giving the order to kill Trevor in a year. Once the hitman receives the email, he will go off the grid and become unreachable until after he kills Trevor.

In this case, the threat is not imminent (unless we adopt a highly counter-intuitive notion of imminence). But the necessity condition, with respect to force used to prevent

---

[118] Rodin, *War and Self-Defense* (Oxford: Oxford University Press, 2002), p. 42.

[119] That imminence is merely a proxy for necessity (and that necessity is what matters when they come apart) is by now a mainstream view among criminal law theorists. This is perhaps the most significant insight emerging from the scholarly debate on the plight of battered women terrorized by abusive husbands. For example, see Jeremy Horder, "Killing the Passive Abuser: A Theoretical Defense," in Stephen Shute and A.P. Simester (eds.) *Criminal Law Theory: Doctrines of the General Part* (Oxford: Oxford University Press, 2001), p. 292; Alafair S. Burke, "Rational Actors, Self-Defense, and Duress: Making Sense, Not Syndromes Out of the Battered Woman," 81 *North Carolina Law Review* 81 (2002): 211-316, p. 271. Cf. Kimberly Kessler Ferzan, "Defending Imminence: From Battered Women to Iraq," *Arizona Law Review* 46 (2004): 213-262, pp. 253-254.

[120] Helen Frowe, *The Ethics of War and Peace* (Abingdon, UK: Routledge, 2015), pp. 78-80.

the sending of the email (e.g., smashing Susan's laptop just before she clicks "send") could nevertheless be satisfied.

The underlying insight here is that the necessity condition enjoins the least harmful course of action but only *ceteris paribus*.[121] Yet sometimes all other things are not equal. Differential prospects of success are relevant in intuitively judging what is permissible. And so the necessity condition straightforwardly requires the less harmful of two options only under the assumption of comparable prospects of success.[122]

Indeed, if the rationale for requiring imminence is cases of imminent danger are ones in which counter-speech is an ineffective remedy, there is no reason to think these are the only such cases. It is plausible that even in some or many cases of dangerous speech that poses a non-imminent danger, such as dangerous speech online, counter-speech remains an ineffective mode of response.[123] What matters is the overall assessment of counter-speech's prospects for success, compared with that of coercion.[124] To insist on counter-speech over coercion even in cases when coercion has a far greater prospect of success unacceptably endangers the prospective victims of the incited harm.

This leads to a second reason why we should be skeptical that the necessity condition requires an imminence condition on restricting dangerous incitement. Even in cases in which both coercion and counter-speech have equal prospects for success, it

---

[121] Here I follow Lazar, "Necessity in Self-Defense and War," pp. 6ff. Further, though there is clearly a fact-relative sense of necessity, prospective necessity judgments (like all judgments) are made by agents within an evidence-relative perspective. And there are limits on how much evidence we can reasonably be expected to gather about the effectiveness of alternative courses of action. In the context of deliberating whether to criminalize some wrongful conduct, we rightly worry about harms that will befall victims while ineffective alternatives are tested endlessly while coercion postponed. Ignoring this point has led some scholars of criminal law to imagine the necessity condition on criminalization to be much more demanding than it is actually is. For example, Husak imagines that the necessity condition on criminalization would "require the state to conduct a series of experiments in which alternative strategies to attain its objective are implemented and found to be deficient" (*Overcriminalization*, p. 158). But this would be an implausibly demanding interpretation of the principle. A suitably constrained necessity condition would not require endless experiments; it would require something closer to what Husak recommends, which is that "alternatives to given [criminal] laws be identified and assessed" (p. 158). A similar, more plausible principle is defended in Andrew Ashworth, *Principles of Criminal Law* (3rd edition) (Oxford: Oxford University Press, 1999), pp. 67-68.

[122] E.g., Lazar, "Necessity in Self-Defense and War," pp. 10ff.

[123] This is a frequently observed point—simply an application of the familiar claim that the "marketplace of ideas" regularly fails to promote the truth. As Larry Alexander aptly notes, "time is but one of several factors that bear on the efficacy of counter-speech Alexander, "Incitement and Freedom of Speech," p. 109. See also Greenawalt, *Speech, Harm, and the Uses of Language,* pp. 116-118, and Barendt, "Incitement to, and Glorification of, Terrorism," p. 458.

[124] The mere fact that agents have the capacity to change their minds in response to counter-speech—as emphasized by Cohen, "Freedom of Expression," p. 233—clearly does not mean they will.

is not enough to show that counter-speech imposes fewer costs on the wrongdoers. The costs that must be borne by others matter, too, which are morally weighted in the calibrations of overall harm. Suppose I can disarm you, with the same likelihood of success, by shooting you lethally from a distance, or by exploiting my knowledge of pressure points during hand-to-hand combat. *Ceteris paribus*, it follows that I should disarm you through hand-to-hand combat. But if *ceteris paribus* does not hold, this need not follow. For example, if hand-to-hand combat will foreseeably result in you blinding me in one eye before I nevertheless hit your pressure point and disarm you, then I am intuitively permitted to shoot and kill you from a distance instead, even though the *overall* amount of harm thereby caused is greater.[125]

So even if coercion and counter-speech were equally efficacious in combatting non-imminently dangerous speech, we must attend to the matter of cost. Suppose that we lived in a just society in which the (just) criminal law was publicized and enforced extremely well. Suppose that, in this society, by criminalizing dangerous incitement, the state could largely eliminate such speech, given the deterrent effects of the criminal prohibition, all at a reasonably low cost to the political community. In contrast, suppose that the counter-speech alternative is extraordinarily demanding. It would require that citizens be constantly vigilant, arguing against any dangerous speech they encounter and even seeking out dangerous speech in order to challenge it, and spending considerable time and energy engaged in the effort in defuse the dangers of dangerous speech.[126] And it would require that the state invest considerably in counter-speech, both by engaging directly in it[127] and by providing institutional support for its occurrence.[128] Is it plausible that dangerous speakers who incite wrongdoing have a moral claim that citizens resort to counter-speech, rather than coercion, if the former involves far greater cost to the political community? The answer seems to me to be no.

Consider those who incite terrorism online through posted videos. Would it violate the necessity condition to delete these videos, thereby suppressing the speech of these individuals? Suppose the alternative is to identify the thousands of impressionable

[125] This also follows Lazar, "Necessity in Self-Defense and War," pp. 6ff.
[126] See my discussion of unreasonable costs in "Terror, Hate, and the Demands of Counter-Speech."
[127] Brettschneider, *When the State Speaks.*
[128] Cohen, "Freedom of Expression," p. 246.

viewers who are sitting in their bedrooms watching these videos, and then subject them—somehow—to counter-argument. Even if this alternative would be as effective as simply banning these websites and deleting the videos—something I doubt[129]—they are clearly costly. Given that the necessity condition is rightly sensitive such costs, it is highly plausible that there will be cases of non-imminently dangerous speech in which coercion will satisfy the necessity condition.

* * *

Even if necessity cannot save *Brandenburg*, it may be that concerns of wide proportionality can. It is easy to see why those who regard the state as inherently corrupt and irremediably untrustworthy—fated to abuse or misunderstand its own moral authority—might endorse a constitutional principle that protected nearly all incitement, simply to guard against abuse. This pessimism strikes me as unwarranted; the idea that considerations of wide proportionality enjoin all societies, everywhere, to adopt a constitutional prohibition on laws restricting dangerous speech, except when violence is imminent, seems implausible. But I don't adjudicate the matter here, since it is largely a complex empirical matter concerning the likely effects of distinct policies within different jurisdictions, with their differential susceptibility to political abuse. The theoretical point remains, which is that it is the analysis of wide proportionality that is at stake, since it concerns the question of whether a moral duty should or should not be enforced due to its effects on non-liable parties.

Whatever the right policy conclusion on this matter, at least one point is clear. We should dispense with the pervasive American myth that those who endanger others by inciting their listeners to harm them have a moral right to do so. They do not; if I am right, they have weighty duties to refrain. While the enforcement of these duties may be impermissible all-things-considered, it is false to think that such enforcement would typically wrong the dangerous speakers themselves.

[129] Cf. Iginio Gagliardone et al, *Countering Online Hate Speech* (New York: UNESCO, 2015). See also Alexander Tsesis, "Hate in Cyberspace: Regulating Hate Speech on the Internet," *San Diego Law Review* 38 (2001): 817-874.