CONTENTS

1

Foundations

ON AUGUST 14, 2024, a woman appeared at a courthouse in Chester, England, to plead guilty to committing a speech crime. In the preceding weeks, riots had engulfed the nation after the murders of three young girls. The match that lit the fire was a rumor: that the attacker was a recently arrived Muslim refugee. As chatter spread across social media platforms, hundreds took to the streets, targeting mosques and hotels where asylum-seeking migrants were known to be housed. Rioters were egged on by an incendiary online chorus; the woman appearing at the courthouse in Chester was one of many. When she saw a photo of citizens gathered to help repair a mosque that was damaged in the riots, she sent the post that sealed her fate: "It's absolutely ridiculous. Don't protect the mosque. Blow the mosque up with the adults in it."[1]

This case and many like it were prosecuted thanks to new state powers conferred by the UK's Online Safety Act, brought into law on October 26, 2023. The law creates a raft of new internet speech crimes, including encouraging or assisting serious self-harm, sending knowingly false information intended to cause harm, and (the crime invoked in this case) sending threatening communications. Most ambitiously, the Act requires social media platforms to combat illegal content on their networks.[2] The crimes deemed "priority offences" under the law—which large platforms have the greatest responsibility to police—are manifold. They encompass uncontroversially prohibited content like child sexual exploitation and abuse material, death threats, illegal weapons sales, and the coordination of human trafficking. They also include a wide variety

of terrorism offenses, including fundraising for proscribed terrorist organizations and weapons training. More controversially, the Act encompasses offenses such as "expressing an opinion or belief supportive of a proscribed organisation," "encouragement of terrorism," and "dissemination of terrorist publications." It also mandates platforms to suppress religious and racially aggravated harassment, and to empower users to avoid content that incites hatred. Given the challenges of accurately deploying artificial intelligence systems to detect such content (as large platforms must), and the stiff penalties for noncompliance, platforms are arguably incentivized by the regulation to err on the side of over-removal rather than under-removal.[3]

Were the Online Safety Act introduced in the United States, much of the law would doubtlessly be struck down by the courts as an unconstitutional affront to freedom of expression. This is, in part, because of the content of the relevant crimes. Our introductory case is illustrative: however repugnant it may be to encourage the bombing of a mosque, such speech is protected under the First Amendment to the US Constitution unless it is both likely and intended to cause imminent harm (determined in the famous 1969 case *Brandenburg v. Ohio*).[4] Likewise, speech advocating terrorism or even praising terrorists is broadly protected speech. Suppressing that speech—whether by directly punishing speakers or ordering platforms to refrain from hosting it—is, on the dominant legal view, constitutionally impermissible. Likewise, laws against hate speech, mis- and dis-information, and self-harm content would also face serious constitutional obstacles in the US. Moreover, even for unprotected speech that could, in principle, be constitutionally suppressed, the laws limiting such speech could still be struck down if they had the collateral side effect of burdening protected speech—for example, by chilling its expression, or incentivizing platforms to engage in overbroad removal practices. For these reasons, it is safe to say that much of the Online Safety Act simply would not pass American constitutional muster. Nor would the even more onerous Digital Services Act (DSA) in the European Union, which places expansive demands on platforms to crack down on illegal content and police their spaces for risks of harm.

The world's democracies are thus in sharp disagreement on the permissibility of limiting speech for the purpose of preventing harms to which it might lead. This divergence raises fundamental questions of political philosophy. The general one concerns when, if ever, it is permissible for the state to suppress speech to prevent harms to which the speech might lead. In many cases, speech will lead to harms through its persuasive effects on listeners' beliefs about what they ought to do. The paradigmatic category of such speech, at the center of this book's inquiry, is *incitement*, in which a speaker encourages listeners to engage in some wrongful harm. (Incendiary speech that receives protection is often termed *advocacy* instead of *incitement*, but I will show this distinction is difficult to ground.) Adjacent and sometimes overlapping to this surprisingly undertheorized category are certain forms of so-called hate speech; varieties of dangerous mis- and dis-information; and speech encouraging self-harm. (While incitement is our story's lead character, these related categories will make repeated appearances in what follows.)

Yet the American approach is an outlier. All democracies have legal protections for freedom of expression—yet that freedom is balanced against a range of other rights, values, and interests. Most democracies restrict more expansive categories of incitement, hate speech, and even disinformation, and have moved to compel online networks to limit the spread of such content. That balancing approach is also dominant in international human rights law, which emphatically recognizes the freedom of speech as a fundamental human right while also affirming specific restrictions on harmful speech. So, Article 19 of the International Covenant on Civil and Political Rights (ICCPR) declares that "everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds"—but then immediately notes that this right "carries with it special duties and responsibilities." The subsequent ICCPR article proceeds to endorse legal restrictions on "advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence," as well as speech constituting "propaganda for war."[5]

Strikingly, this globally dominant position is repeatedly condemned as both illiberal and antidemocratic by luminaries in the philosophical

literature. The notion that various forms of dangerous speech morally ought to receive capacious legal protection has been endorsed by some of the most important liberal political philosophers of the twentieth century.[6] Such views constitute the dominant position most commonly and closely associated with the political philosophy of liberalism. On this family of views, respecting people as autonomous moral agents requires permitting them to say what they think, and hear what others think. That is especially so considering their status as democratic citizens, who must be unconstrained in their prerogative to debate a wide range of normative and empirical propositions. While there are many versions of this argument—corresponding to different accounts of what precisely justifies the moral principle of freedom of expression—it constitutes the central philosophical strategy in defense of extensive legal protections for dangerous speech.

The Book's Aim

Who is right? The answer to that question is more complicated and contingent than partisans of the free speech culture wars tend to suggest. My view is that both sides of this debate have captured crucial insights, which a proper philosophical theory of free speech and its limits must integrate. Free speech *is* among our most important rights. We appropriately revere legal protections such as the First Amendment as bedrock to our democratic society and civic identity. The principle of viewpoint neutrality, whereby the state should refrain from restricting disfavored viewpoints, is fundamental. And given the history of unjust censorship by oppressive states, everyone should be highly skeptical of efforts by governments to interfere in their population's communications.

Yet this laudable skepticism has hardened into a rigid orthodoxy in political and legal theory, whose proponents claim too much at the level of fundamental principle. Against this orthodoxy, I argue that our moral right to free speech, while vital, is nevertheless constrained by enforceable moral duties we owe to others to refrain from certain forms of wrongfully harmful speech. Further, I argue that intermediaries, such

as social media platforms, have enforceable moral duties to restrict such speech. Laws enforcing these duties are not inherently forbidden by political morality. While such laws are presumptively suspect under the doctrine of viewpoint neutrality, I believe that presumption is easier to rebut, in principle, than commonly supposed. Those concerned about the dangers posed by incitement and other forms of harmful speech seem to have grasped this basic insight, even if they have not persuasively articulated it. Much of the first half of this book seeks to spell out and defend this position.

Still, it hardly follows that those keen to restrict dangerous speech are correct all-things-considered. Even if a speaker lacks a moral right to engage in certain speech and has a moral duty to refrain from that speech, it doesn't follow that her moral duty should be enforced. This book's further insight, then, is that our fundamental moral rights and duties, which define the deep morality of a particular normative domain, should not be straightforwardly codified into legal rights and duties. A wide range of morally inflected pragmatic considerations properly enter the picture when determining the scope of legal rights, and especially the nuanced contours of public policy. There are, I will argue, innumerable ways in which morality and justified policy might come apart—to put the point roughly, in which something might be justified in theory but not in practice. Such an incongruence is manifest in many contexts; a central theme of just war theory over the past two decades is the divergence between the morality and justified law of war. My contention is that the same divergence plausibly arises in the context of free speech. This book thus attempts to do for free speech theory what recent scholarship has accomplished for just war theory.[7]

Why might the ethics of speech, and the justified law of speech, come apart? One reason is that, to justify a speech restriction, it is not nearly enough to show that the speech is harmful (though this is where most of the prohibitionists' focus lies). Nor is it enough to show (as I shall) that certain harmful speech violates a duty of justice that citizens owe to one another. Instead we must show that the enforcement of the relevant duties would be necessary and proportionate. I argue that many familiar claims in the free speech literature, often presented as bearing

on matters of fundamental principle, are better construed as bearing on the all-things-considered permissibility of enforcing our speech-related duties. Many common concerns—about the chilling effects of legislation, or the precedents for future abuse, or about the potential superiority of counter-speech as a remedy, or about the ineffectiveness of speech restrictions—all enter the analysis at this stage. The real philosophical debate to be had about wrongfully harmful speech (such as incitement) is not whether it is intrinsically morally protected as free speech (it isn't). It is about whether laws restricting such speech, either by coercing speakers or coercing platforms, can pass appropriately specified tests of necessity and proportionality. Such tests turn on a complex combination of normative and empirical judgments.

By exploring what it would take for laws restricting incitement to pass these tests, my goal is to equip readers with the theoretical framework they need to work through whether particular speech restrictions, in particular jurisdictions, are justified. While I identify the relevant considerations that bear on such judgments, their magnitude is variable depending on the particular law in question, the particular government enforcing the law, and the particular social context in which the law is to be enforced. In this way, I contend, free speech policy is just like other morally significant areas of public policy. Policy that can be justified in one place and time might not be justified at another. The insights I trace will be, I hope, useful to citizens, legislators, and even judges seeking to develop and apply extant legal doctrines, especially concerning ideas of necessity and proportionality.

By defending these insights, the book seeks to achieve three broad goals. First, I offer a systematic reorganization of the philosophical debate on free speech, distinguishing stages of analysis that have been unhelpfully run together. The initial stage concerns the moral duties of speakers to refrain from various forms of harmful communications. The next concerns whether these duties are, in principle, enforceable, or whether the moral right to free speech blocks their enforcement (as many philosophers believe). And the final stage concerns whether the legal enforcement of those duties would be justified all-things-considered. A failure to distinguish these stages has led to confusion in

both public and scholarly debate. For example, pragmatic concerns relevant to the proper specification of the law are irrelevant to the foundational issue of who has a moral right to say what, and often turn on contingent empirical rather than normative claims.

Second, I advance the debate through a series of novel substantive arguments. While a prevailing orthodoxy allows restrictions on speech when the harm threatened is *imminent*, I contend that this temporal requirement is misguided (especially in the era of social media). Once we see why it is acceptable to restrict speech in the case of imminent harm—because the speaker is violating an enforceable moral duty he owes to others not to endanger them—we can recognize a wider range of cases in which restrictions can, in principle, be justified. I defend the moral permissibility of greater restrictions on speech than many liberals are prepared to allow, while still insisting that countless restrictions on speech around the world are unjust (including many restrictions in democratic states). I also develop and defend a philosophical analysis of the moral duties of social media platforms and these duties' permissible enforcement.

Finally, through these arguments I seek to demonstrate that concerns about freedom of speech are not part of some sui generis silo of normative inquiry, but rather are continuous with wider issues across political, moral, and legal philosophy. For example, I show that speakers' duties to refrain from incitement flow from broader moral obligations not to risk imposing undue harm on others; that arguments about the democratic role of free speech cannot be adjudicated without delving deeply into democratic theory; and that questions about the all-things-considered permissibility of speech restrictions require engagement with myriad insights from the ethics of defensive harm, criminal law, and human rights law.

The theory that emerges serves to undermine the false dichotomy that has come to define the public debate on this topic, whereby either we permit harmful speech or we ban it. In so doing, it reveals the debate on the limits of free speech for what it is: a complicated policy matter hinging on a wide range of contestable normative and empirical considerations, emphatically within the domain of reasonable disagreement.

The public discourse of all liberal democracies would benefit from seeing this controversy as a nuanced and difficult one, rather than the feverish culture war it has become. Further, we should accept a wider range of positions on free speech as legitimate within a liberal democracy, rather than insisting, implausibly, that one's preferred view is the only reasonable option.

Having identified my topic and previewed the arguments to come, I'll spend the rest of this chapter setting the stage for the rest of the book. First, since my overarching topic is freedom of speech and its limits, I will disaggregate the variety of phenomena to which "freedom of speech" can refer in both popular and scholarly discourse, clarifying which are the focus of this book. Next, I will discuss the specific sense of free speech that dominates my initial set of arguments: a natural moral right to freedom of expression. I will then defend a specific conception of this moral right: the *human powers theory*. On this view, the moral right to free speech is essential to the development and exercise of two distinctive human capacities: the capacity to be a self-determining author of our own life, and our capacity as a moral agent (which chiefly involves recognizing and respecting others' capacities to be self-authors). This is at once a capacious and uncontroversial theory of free speech, which ties together a variety of extant views in the scholarly literature, restating their various commitments as part of a single doctrine that can constitute an ecumenical starting point for the more controversial arguments to come. Setting out a demanding theory of the moral right of free speech is an essential starting point of the book, since if there were no such right, restrictions on speech would be substantially easier to justify. Finally, I will briefly reflect on the book's methodology and then recapitulate the precise arguments to come through a conspectus of the chapters ahead.

What Is Free Speech?

In the philosophical literature, the terms *freedom of speech*, *free speech*, *freedom of expression*, and *freedom of communication* are mostly used equivalently.[8] I shall follow that convention here, notwithstanding the

fact that these formulations evoke subtly different phenomena. For example, it is widely understood that artistic expressions, such as dancing and painting, fall within the ambit of this freedom, even though they don't straightforwardly seem to qualify as *speech*, which intuitively connotes some kind of linguistic utterance or writing. Still, they plainly can qualify as communicative activity, conveying some kind of message, however vague or open to interpretation it may be.[9]

Yet the extension of *free speech* is not fruitfully specified through conceptual analysis alone. The quest to distinguish speech from conduct, for the purpose of excluding the latter from protection, is notoriously thorny, despite some notable attempts (which I will have occasion to discuss in chapter 3). As John Hart Ely writes concerning the burning of draft cards to protest the Vietnam War, such activity is "100% action and 100% expression."[10] It is only once we understand *why* we should care about free speech in the first place—the values it instantiates or serves—that we can evaluate whether a law banning the burning of draft cards (or anything) violates free speech. It is the task of a normative *conception* of free speech to offer an account of the values at stake, which in turn can illuminate the kinds of activities wherein those values are realized, and the kinds of restrictions that manifest hostility to those values.[11] For example, if free speech is justified by the value of respecting citizens' interest in hearing many points of view and to make up their own minds, then banning the burning of draft cards to limit the viewpoints to which citizens will be exposed is manifestly incompatible with that purpose. (If, in contrast, such activity is banned as part of a generally applied ordinance restricting fires in public, it would likely raise no free speech concerns.) I turn to the substantive issue of justifying free speech in the next section. Before doing so, let me spell out some further distinctions in how the terminology of free speech is deployed.

First, as already intimated, we can distinguish between the *morality* of free speech and the *law* of free speech. In political philosophy, a standard approach is to theorize free speech as a requirement of morality, tracing the implications of such a theory for law and policy. Note that while this is the order of justification, it need not be the order of

investigation; it is perfectly sensible to begin by studying an existing legal protection for speech (such as the First Amendment in the US) and then asking what could justify such a protection (or something like it).

But morality and law can diverge. The most obvious way they can diverge is when the law is unjust. Existing legal protections for speech, embodied in the positive law of particular jurisdictions, may be misguided in various ways. In other words, a justified legal right to free speech, and the actual legal right to free speech in the positive law of a particular jurisdiction, can come apart. In some cases, positive legal rights might protect too little speech. For example, some jurisdictions' speech laws make exceptions for blasphemy, such that criminalizing blasphemy does not breach the legal right to free speech within that legal system. But clearly one could argue that a justified legal right to free speech would not include any such exception. In other cases, positive legal rights might perhaps protect too much speech. Consider the fact that, as a matter of US constitutional precedent, the First Amendment is held to protect much of the hate speech banned elsewhere. We could agree that this is so as a matter of positive law while disagreeing about whether it ought to be so.[12] (This book is centrally concerned with the right to free speech as a matter of political philosophy, tracing the normative implications for downstream questions of constitutional and legislative decision-making.)

Second, we can distinguish rights-based theories of free speech from non–rights-based theories. For many liberals, the legal right to free speech is justified by appealing to an underlying moral right to free speech, understood as a natural right held by all persons. (Some use the term *human right* equivalently, though the appropriate usage of that term is contested.)[13] The operative notion of a moral right here is that of a *claim-right* (an idea developed in Hohfeld's influential analysis);[14] it thereby correlates to moral duties held by others (paradigmatically, the state) to respect or protect the right.[15] Accordingly, this right correlates to a directed duty, a duty owed to the right-holder, such that its violation wrongs her. As Matthew Kramer helpfully elaborates, "a

claim-right confers deontic protection on its holder against the interference or uncooperativeness of some other person(s)."[16] So, a claim-right to free speech confers a *freedom* to communicate, where that freedom is understood as a deontic protection against the interference or uncooperativeness of others, including states.[17]

What exactly this requires normatively—what, in other words, this is a claim-right *to*, determining what counts as a violation—will again depend on what justifies free speech (to which I turn momentarily). Across all theories, paradigmatic violations of free speech include the effort to stop people from expressing or hearing particular disfavored points of view, or forbidding discussion of certain topics (respectively termed *viewpoint-based* and *content-based discrimination* in the lexicon of American constitutional law).[18] These violations are typically understood in terms of the government's purpose in acting, thereby condemning efforts to restrict viewpoints or content whether or not they succeed.[19] Violations may also include certain unintended collateral effects of otherwise justified speech regulations, such as laws that have substantial chilling effects on protected speech at the borderline of the speech they target (incentivizing audiences to self-censor). Depending on one's theory, violations may also involve failures to discharge positive obligations to provide people with opportunities to speak to and hear from others, or even to develop their capacities as speakers and listeners.

Where the moral right to free speech is instantiated in law as a legal right, that moral protection will be codified as a protection within the legal system. But crucially, even where no such legal protection obtains, the moral protection binds nonetheless. Such a right is natural in that it exerts normative force independently of whether anyone thinks it does, and regardless of whether it is codified into the law. A tyrannical state that imprisons dissidents acts unjustly, violating moral rights even if there is no legal right to freedom of expression in its legal system.

For others, the underlying moral justification for free speech law need not come in the form of a natural moral right. For example, some consequentialists (that is, those who think morality is strictly a matter

of promoting good consequences) might favor a legal right to free speech on the grounds that it promotes the general welfare, without thinking that it tracks any underlying natural right. Or consider democratic theorists who have defended legal protections for free speech as central to democracy. Such theorists may think there is an underlying natural moral right to free speech, but they need not (especially if they hold an instrumental justification for democracy). Or consider theorists who have argued that free speech functions as a deontological side-constraint on legitimate state action, requiring that the state always justify its decisions in a manner that respects citizens' autonomy.[20] This theory does not cast free speech as a right, but rather as a principle that forbids the creation of laws that restrict speech on certain grounds. In the Hohfeldian analysis, such a principle may be understood as an *immunity* rather than a claim-right.[21] Finally, some "minimalists" (to use a designation from Joshua Cohen) favor legal protection for speech principally in response to government malice, corruption, and incompetence.[22] Such theorists need not embrace any fundamental moral right, either.

Third, among those who *do* ground free speech in a natural moral right, there is scope for disagreement about how tightly the law should mirror that right (as with any right[23]). It is an open question what the precise legal codification of the moral right to free speech should involve. A justified legal right to freedom of speech may not mirror the precise contours of the natural moral right to freedom of speech. As I will show, a raft of instrumental concerns enters the downstream analysis of what any justified legal right should look like; hence a defensible legal right to free speech may protect more speech (or less speech) than the underlying moral right that justifies it. For example, even if the moral right to free speech does not protect much hate speech, such speech may still merit *legal* protection in the final analysis (say, because it would be too risky to entrust states with the power to limit those communications).[24]

Importantly, while a legal right to free speech might be embedded in some constitutional text, it can also be created statutorily. Likewise, one could imagine protecting the core of the legal right constitutionally, but

expanding the right through additional statutory protections that are created through ordinary legislative action. Further, note that when international lawyers talk about the *human* right to free speech, they are sometimes referring not to the natural moral right, but rather to the legal right (either as it exists positively, or normatively should exist, within international human rights law).

Finally, free speech might be invoked to characterize a particular *cultural ethos*, quite apart from the issue of moral and legal rights. This might be merely a positive description, as in "universities tend to have a cultural ethos of free speech," leaving it open whether that is in fact a good thing. Such a cultural ethos could also be normatively defended, of course, as desirable or justified. Consider the fact that social media companies routinely express support for free speech as a basic principle. This might be explained by the view that not only governments but also powerful private actors have duties to respect free speech (a point to which I shall return in chapter 4). But it might also be explained by the notion of a free speech ethos. As Robert Simpson aptly notes, the idea of a free speech culture is one "in which all are encouraged to speak their minds and to work through their disagreements in debate and discussion, instead of trying to silence or ostracize opponents."[25] The claim that people are too quick to "cancel" their political opponents—ostracizing them or getting them fired—and that this is in tension with free speech—is likewise best construed as invoking the idea of a free speech ethos. As with free speech as a principle, an ethos of free speech—for whatever reason—might protect more speech than speakers have a moral claim-right to engage in.

Unless we distinguish between different senses of the concept, we risk talking past one another. I have pinpointed the meanings of the concept that will loom largest in the discussion ahead. While this framing is not intended to be exhaustive, and other framings are plausible, the one I set out is especially perspicuous for making and grasping the arguments to come. In this book, I will be principally concerned with the natural moral right to free speech and its implications for a justified legal right to free speech, though in the course of analyzing such issues I will study various proposed principles of free speech. As this is a work

of political philosophy rather than law, I will not focus on analyzing existing legal protections for free speech and their judicial interpretation, which vary across polities. Even so, the abstract philosophical questions I seek to answer have normative implications for how we should interpret or reform existing speech laws, and my analysis will be useful to lawyers and judges focused on the particulars of free speech law in specific jurisdictions.

The focus on free speech *rights*, rather than principles, has a substantive rationale. A distinctive purpose of rights-talk is to help us see *who is wronged* by any given violation. As Rowan Cruft puts it, a claim-right is "constituted by duties owed to the right-holder, duties whose violation wrongs the right-holder."[26] As noted above, the right to free speech correlates to a directed duty, a duty owed to the right-holder, such that its violation wrongs her. And while it may be that certain speech protections are structured in terms of immunities rather than claim-rights, one important *justification* of immunities appeals directly to their role in protecting claim-rights.[27] By abandoning rights-talk, we risk losing track of who is wronged by violations, and why.

Two final preliminary points about moral claim-rights. First, while many acts are *presumptively* covered by a given right, whether any given act is protected by the right in a given case will depend on the absence of sufficiently weighty countervailing considerations. All of our rights work like this. For example, the right to privacy plausibly protects me against police searches of my vehicle, but during a car-bomb threat in my neighborhood, my privacy interests appropriately yield to security concerns. Likewise, the right to religious liberty presumptively protects a wide array of conduct connected to the practice of one's religion, but only presumptively; whether a given act is in fact protected again depends on an assessment of countervailing costs. Or consider: the rights to freedom of movement and association presumptively permit us to visit large groups of people in their homes; yet during a pandemic, the contours of this right adjust. If this model is correct, then our rights to freedom of movement and association were always best understood as conditional on not posing a danger of transmitting a virus to others, and hence aren't even infringed by (justified) lockdown restrictions.[28]

Accordingly, the right to freedom of speech might protect certain speech in one context, but not in others.[29]

Second, that a natural moral right to free speech, like any claim-right, will always be defeasible (that is, overridable). In an some emergency cases, the right continues to exert its full normative force, but nevertheless may be permissibly infringed, pro tanto wronging the right-holder but with an all-things-considered justification (as when killing one innocent person as an unintended but unavoidable side effect of saving many lives).[30] I doubt that such "lesser evil" cases arise often in the public policy of a well-ordered liberal democracy, where anticipatory steps can be taken to prevent their occurrence, but they are certainly possible.

What Justifies Free Speech?

Human beings have significant interests in communicating to others, and in listening to what others have to say. These interests make it difficult to justify coercive restrictions on people's communications, grounding a moral right to speak (and listen) to others that is properly protected in law. What are these interests?[31]

Theorists have illuminated a variety of considerations that could justify the right to free speech. My view is that we need not choose among them; a plurality of interests together ground the right. Still, these interests can be placed within a unifying theoretical framework—an organizing conception of the right to freedom of speech—which explains why they belong together. To my mind, this conception is most convincingly embedded in a broader *liberal* political philosophy, centered on the ideal of citizens as free and equal. For the avoidance of doubt, "liberal" here refers to the family of political theories grounded in the idea that all human beings are of immense and equal value, and that the fundamental way of respecting that value is to accord people the freedom to chart their own paths in life, consistent with the rights of others to do the same. So understood, liberals need not self-identify as "progressive" in the American sense, nor do they need to self-identify as "neoliberal" in the sense of promoting free

market economics—though of course many progressives and neoliberals will count as liberal, as will many conservatives. The American Declaration of Independence and Bill of Rights to the US Constitution, the European Declaration of Human Rights, the Charter of Fundamental Rights of the European Union, the UK Human Rights Act, the International Covenant on Civil and Political Rights, the South African Constitution, the Indian Constitution—these (just to name some examples) all articulate liberal ideals in my sense.

On my own favored liberal view, powerfully defended in the twentieth century by John Rawls, respecting others as free and equal involves respecting two core human powers. The first I will call *self-authorship*—a capacity to frame, revise, and pursue a conception of a good and meaningful human life. The second I will call *moral agency*—a capacity to understand and act on the moral duties we owe to each other, particularly duties of justice. On this view, our fundamental interests in the development and exercise of these two human powers properly ground and guide the specification of our basic liberties and the terms of social cooperation—including our conception of free speech.[32]

Why does freedom of speech enable the full development and exercise of our human powers of self-authorship and moral agency? The answer is already at hand; the extant theories of freedom of speech are each fruitfully understood, I submit, as identifying an important dimension along which freedom of speech serves the development and exercise of those powers. I will now detail how I map out this normative space. But this space can be mapped in different ways, and I don't think there is any one perfect way to tell this story.[33] In telling my variation of it, I draw heavily on others' work. While Rawls himself dedicated relatively little attention to freedom of speech,[34] other liberals in the Rawlsian tradition certainly have. So, many of the considerations covered in Joshua Cohen's touchstone essay "Freedom of Expression" appear below. I am also substantially indebted to Seana Shiffrin's groundbreaking thinker-based theory, which explicitly stresses the fundamental role of free speech for the development and exercise of our moral agency, also in an implicitly Rawlsian register. The account I posit here seeks to assimilate a wide range of insights on freedom of expression in the

history of political thought and contemporary political and legal theory. John Stuart Mill's famous views on the epistemic value of free speech, Alexander Meiklejohn's on the importance of free speech for democracy, Rae Langton's on the significance of speech-acts for free speech—these are just some of the many views I seek to encompass. The goal is to stipulate a reasonably comprehensive but fairly ecumenical liberal vision for why free speech should matter to us—one I shall label the *human powers theory*. Having set the stage with such a view, we can then spend the book inquiring into its limits, and how to think about them. (Liberals who eschew the Rawlsian apparatus can still help themselves to the basic ideas that follow; the specific interests I pinpoint will resonate, I hope, for non-Rawlsians, and even some non-liberals.)

Self-Authorship

Start with the most vaunted human capacity in the liberal tradition: the capacity to be the author of one's own life plan. This is the capacity to frame, pursue, and revise one's *conception of the good*—a rational plan setting out various projects and goals for one's life and how to attain them. Such a conception may be embedded in a broader comprehensive doctrine, such as an overarching religious faith, but it need not. For many, one's conception of the good will take the form of a pluralistic list. The list may encompass familiar commitments like developing emotionally deep relationships with other human beings, pursuing meaningful work, engaging the human senses and imagination, playing, contributing to one's community, or being a just or good person. For a life to be self-authored, agents must enjoy wide latitude to decide if and how to develop these various capabilities.[35]

Clearly many basic liberties—including freedom of association, freedom of religion, and freedom of occupational choice—support the development and exercise of the human power of self-authorship. How does freedom of communication distinctively do so?

First, communication is fundamental to processes of *education*, broadly conceived to encompass our upbringing and socialization. Our developmental process is essentially communicatively mediated,

enabling us to become the kinds of beings who can form a map of our world (physical and social)—without which navigation toward the good would be inordinately difficult. Such processes enable us to develop basic skills in theoretical and practical reasoning.[36]

Second, we have interests in learning *information* relevant to selecting and pursuing a conception of the good.[37] For example, we have interests in learning about the wide variety of world religions, and criticisms of those religions; in learning about the wide variety of occupations one can pursue in one's society; and in learning about what is happening in the world so that we can decide how best to execute our plans. Because we cannot learn all relevant information directly ourselves, we must communicate with others to glean what they have learned.

Third, we have interests engaging in *deliberation* about what a good life might involve, as part of a process of working out what one really thinks. As Shiffrin stresses, we often need to externalize and talk through an idea before we can figure out what we really think.[38] We decide what to do, in part, by discussing proposed projects and goals with others, so that we can hear their reactions and feedback.[39]

Fourth, speech enables us to exercise our *imagination*; we use speech to fantasize and construct hypothetical scenarios, in order to envisage them fully and with sufficient detail.[40] For many, the exercise of imagination is also partly constitutive of a good life (rather than merely an instrumental means to an end), and is often bound up with our interests in play.

Fifth, communication is indispensable for flourishing *relationships* with other human beings. Some of these relationships, such as those with our families and friends, will be among the most significant of our lives, and central to what we take to be good and valuable about our time on Earth. The same applies for our relationships with those with whom we engage in other cooperative activities, including recreational ones. Even the fleeting conversations we have with those from whom we order our coffees are an important source of connection that brings benefits to our life.[41]

Sixth, communication enables our interests in *self-disclosure*: in telling others who we are and what we think. We have an interest in telling others

what we think, in part, because we have an interest in being recognized for who we are—an interest in the affirmation of our unique identity.[42] Such expression may also reap cathartic benefits, even if it has no likely impact on how audiences see us. There is value in shouting who we are from the rooftops, as it were, in serving as a witness to our own convictions.[43]

Seventh, communication serves our interests in *understanding others*.[44] As Joseph Raz notes, this may be because our conception of the good is inherently social, and we want to find out who shares the beliefs we already possess.[45] In other cases, knowing what others think is valuable for strategic reasons—in evaluating who is likely to be an ally or an obstacle in the pursuit of one's goals. People might be able to guess our beliefs without our telling them, but only speech can convey accurately and precisely our thoughts.[46]

Eighth, communication serves our interests in *persuasion*. We have interests in seeking to convince others to join us in some task, or even fully adopt our own conception of the good.[47] We may seek to do that, because our conception itself enjoins proselytism. Or we might do so for instrumental reasons, because it serves us if others act in some way (which can be as ordinary as persuading people to join our sports team or choir).

Ninth, communication serves our interests in *coordination and mobilization*. Communication is often required to coordinate successfully on a joint activity. Many of the meaningful projects in our lives will involve coordination. Note that we often pursue persuasion and coordination together—trying to inspire people to act together for certain purposes. In this way, communication is essential to mobilization.

Tenth, communication serves our interest in *independent judgment*. We have an interest in making up our own minds about what to think about the huge variety of questions pertinent to framing, revising, and pursuing a life plan.[48] Such an interest is plausibly compatible with some deference to experts. But such deference, to be compatible with this interest, should be based on a higher-order autonomous judgment that the relevant putative epistemic authority is, in fact, an epistemic authority.

This exposition has sought to pinpoint ten overlapping interests connected to the development and exercise of our self-authorship power,

indicating how communication serves those interests.[49] To be sure, exercising our conceptions of the good sometimes requires communicative activity that may not cleanly fall into one of these areas. For example, our religious faith might require us to engage in prayer or worship—either self-directed, directed to our deity or deities, or directed to others. Yet restrictions on such worship, while running afoul of religious freedom, would also decidedly run afoul of freedom of speech.

I have aspired to specify these interests in ways that resonate with those who do not self-identify as Rawlsian liberals, or even as liberals. Yet when divorced from a liberal view, their normative significance becomes less clear. Suppose one held an illiberal perfectionist view—according to which the good life consisted in a particular path, and our moral obligations to one another involved ensuring, through coercion if necessary, that we walked down that path. For such a view, freedom of speech would lose much of its significance. In fact, for illiberal perfectionists, freedom of speech becomes dangerous, enabling people to convince others to embrace false or unworthy conceptions of the good. The significance of freedom of speech arises precisely because the capacity to frame and pursue a conception of the good is one that agents are *entitled* to exercise freely—so that they can come to their own conclusions about what life projects are worth pursuing. As C. Edwin Baker argues, a central purpose of free speech protections is to "delineate a realm of liberty for self-determined processes of self-realization."[50] A core role of the liberal state is to provide the social conditions for the development and exercise of our self-authorship powers, not to direct us down one path. By respecting free speech, it partly fulfills that duty.[51] If free speech were untethered from an underlying liberal political philosophy, its justification would be far more tenuous.[52]

The Human Power of Moral Agency

Now let me turn to the development and exercise of our other distinctive human power: *moral agency*.[53] *Moral agency* refers expansively to our capacities to reflect on and deliberate about what we owe to others, to come to conclusions about what we have duties to do, and to guide

our conduct effectively in accordance with those conclusions. By enabling our moral agency, freedom of communication is an *enabling right*: a right needed to discharge duties that we owe to one another.[54]

Such agency also involves the capacity to understand and evaluate prevailing moral norms and conventions, as well as to understand and evaluate the legal rules of any given jurisdiction. Moral agency also encompasses *political agency*. In a democratic society, political agency names our capacities to participate in democratic politics (by voting, seeking office, or otherwise contributing) and thereby discharge our proper role as citizens. This role includes *advocating* for legal and social change and working to persuade and mobilize others in its service. Such advocacy may take the form of *advocating for others*, and it may take the form of *self-advocacy*. Political agency is obviously also valuable in nondemocracies, even if it more circumscribed. (Many autocracies hold public consultations, and all are sensitive to public opinion; speech is therefore immensely valuable in such jurisdictions, even if it receives lamentably less protection.) Finally, moral agency can also take the form of *exercising one's normative powers*, such as *promising*, *consenting*, and *waiving* (and their various legal analogues).

Our human power of moral agency clearly supports our self-authorship power in various ways; they are mutually supportive and interdependent. For example, exercising our normative powers (such as waiving another's duty not to touch us) is central to pursuing a wide variety of valuable activities and relationships. Further, advocating for oneself in the political space, and winning important legislative achievements, clearly expands the space within which one can engage in self-authorship. (Think of groups campaigning for greater religious freedom.) Finally, for many, working on behalf of other human beings, as part of a struggle for greater justice, is itself understood to be part of the good life.

For these reasons, it is difficult neatly to disentangle our interests in the development and exercise of our human powers. Unsurprisingly, then, most of the above reflections on how communication serves self-authorship apply, mutatis mutandis, to moral agency. It is worth clarifying explicitly how. Here, then, is the list of moral agency interests[55] served by communication:

First, communication is indispensable to the developmental process of *moral education*, both informal and formal. We learn our first lessons about right from wrong through communications with our parents and teachers, and we subject those lessons to further scrutiny over the course of our development—through childhood and into adulthood. This involves, at an early stage, learning how to understand and follow rules, all the way up to learning how to criticize and argue about what the rules ought to be.[56]

Second, we have interests in learning *information* relevant to effectively practicing moral agency in one's society. Most grandly, this may involve the rudiments of history, economics, and philosophy. More mundanely but just as importantly, it encompasses information about what is currently happening in one's society and in other societies; who is running for office on what platform; when elections are being held; and other information relevant to exercise one's agency as a democratic citizen.[57]

Third, we have interests in *deliberating* what justice requires. In a democracy, such processes of deliberation between citizens (and public officials) are fundamental.[58] Participation in such deliberation with others is arguably noninstrumentally valuable; but it is also instrumentally valuable to the extent that it better enables us to identify what justice requires and what we ought to do all-things-considered. While ongoing discussion is not sufficient to lead us to the truth on matters of justice, it is nevertheless conducive to that aim.

Fourth, we have interests in the *imagination* of new possibilities, and new futures we might create, as part of our efforts to determine what a just society should involve. Imagination is also vital in our pursuit of innovative solutions to problems we face in public policymaking.

Fifth, we need communication to develop and maintain flourishing political *relationships* with other human beings, especially our fellow citizens. Political and civil society associations are central mechanisms by which political proposals are incubated and advocated. In this way, free communication is itself integral to free association. Further, communication is integral to the ways in which we maintain relations as *equals* with our fellow citizens.[59]

Sixth, we have interests as persons in *independent judgment*—making up our own minds about what to think about questions of justice and morality more broadly. It is presumptively incompatible with that interest to grant others (especially the state) the power to pre-screen what viewpoints we are allowed to entertain.

Seventh, we have interests in *self-disclosure*—telling others who we are and what we think about matters of political and moral conviction. Even if doing so will not make any difference to what people will do, we have an interest in standing up and declaring what we believe is morally right.

Eighth, we have interests in *understanding others*. This is, in part, because successfully exercising our moral agency involves duties to be responsive to what other people believe—paradigmatically, but not only, when others are exercising their own normative powers.[60] In other cases, respect for others involves a willingness to listen to them. In democratic politics, we have especial interests in understanding what others think because public opinion plays an important role in constraining what is feasible.

Ninth, moral agents require communication to engage in *persuasion* (or, if you prefer, *advocacy*), enjoining others to do what the speaker sees as right. Sometimes such persuasion will be morally compulsory.[61] Such advocacy need not merely aim to externalize one's thoughts; it aims to change others' minds with the goal of influencing their decisions.[62] Persuasion, then, is a *speech-act*, with a distinctive locutionary meaning, illocutionary force, and potential downstream perlocutionary effects.[63] That it is a speech-act makes it no less valuable.[64]

Tenth, communication is often required to *coordinate* successfully on a joint activity. Justice requires that we act together for certain purposes to discharge moral duties we owe to each other. Communication is necessary in order to organize politically and pursue the complex coordination that enables us to fulfill our moral duties to one another.[65] In this way, communication is essential to *mobilization*—a core aspect of political activity. Citizens need to be free to mobilize to campaign for their preferred candidates in elections, get out the vote, and organize grassroots campaigns for social change. All of the great

political campaigns in history required active communication between participants.

Eleventh, we have a further interest in *moral accountability*, a distinctively communicative practice involving blame, apology, and so on. Free speech is also needed for *political accountability*, holding our elected officials accountable, and openly discussing and debating the wisdom of domestic and foreign policy.

Finally, we all have interests in *exercising normative powers*, and in many cases such powers are essential to discharging our moral duties. For example, politicians often have duties to make promises to their constituents. Voting is also the exercise of a normative power.[66]

I have reviewed our various interests in exercising our moral agency. As with the self-authorship interests, these moral agency interests are advanced as part of a liberal view. One can imagine an ardent campaigner doubting free speech simply on the grounds that he has already identified the right view about justice, and so wants to restrict communicative liberty so that widespread adherence to his conception can be better achieved and maintained. Of course, one might argue that he *needs* free speech to have a good opportunity to install his preferred theory of justice in citizens' minds. But once he has done that, what argument should dissuade him from curtailing free speech henceforth?

Here it is vital that the defense of free speech is not merely instrumental, but rather grounded in respect and concern for citizens[67] as free and equal bearers of the two human powers. Within this broader framework, the point isn't about ensuring, however one can, that citizens arrive at the best view about justice; it is instead a matter of respecting, and enabling, citizens' exercise of their own judgment about what justice requires. To be sure, citizens might come to the wrong views, and depending how bad those views are, it might be right to stop them from implementing or acting on those views in ways that endanger others—a point to which I shall return in depth. Yet to decide at the outset that one has arrived at the truth about justice, and shut down all further deliberation about the matter, is simply incompatible with respect for the moral capacities of citizens—incompatible with their interest in

independent judgment. As Scanlon influentially puts it, "a person must see himself as sovereign in deciding what to believe and in weighing competing reasons for action."[68] People must be free to decide what to think for themselves. Shutting down communicative channels to subvert or even steer that process is presumptively incompatible with that liberal entitlement. Hence the doctrine of *viewpoint neutrality* in free speech jurisprudence, whereby the state must presumptively refrain from suppressing disfavored viewpoints.

I have completed my brief statement of what justifies freedom of speech as a moral claim-right,[69] focusing on the indispensable role of communication for the development and exercise of our human powers. Many are skeptical that a natural moral right to free speech even exists.[70] If that were my starting point, it would make it enormously easy to justify restrictions or limits on speech, since there is no moral right to free speech to render those restrictions presumptively wrongful. I make the task ahead substantially harder by positing the existence of a weighty natural free speech right.

Methodology

This book follows a standard methodology of free speech theory, and political philosophy generally. On this view, we first test whether a given category of conduct (in our case, speech) is protected by a right by working through the interests, values, and reasons that justify that right.[71] (I will use the term *interests* as encompassing all of these.) Here the initial question is whether the activity engages the core interests that the right exists to respect or protect; the weaker the engagement of those interests by the activity, the weaker the case for the activity's protection under the right. And if certain interests are not engaged at all by some activity, that activity lacks protection under the moral right. (It may still receive protection under the corresponding *legal* right—but that depends on further, instrumental considerations, to be pursued later in this book.)

Supposing that the activity *does* engage the interests, it does not automatically follow that the activity is protected by the right. That

depends on the countervailing reasons that militate in favor of *excluding* the activity from the protection of the right. (In the case of harmful speech, those interests just are the interests of those at risk of harm from the speech.) In some cases, the activity in question might substantially engage the relevant interests; but the countervailing reasons are simply weightier than those generated by the interests. In this way, in specifying a theory of moral rights, the contours of each are adjusted to "make room," as it were, for each other. As Scanlon puts it:

> Rights are rarely absolute in one sense of that term: the norms that a defined right specifies generally do not rule out all interference with the interests that the right protects. In order to protect these interests at tolerable costs, rights need to incorporate exceptions that take account of other interests.[72]

Thus even in cases where we judge that free speech interests are served by some activity, that doesn't necessarily settle the issue. We also need to attend to other interests. This leaves open, of course, how precisely to weigh the relevant interests against one another. Following Scanlon's contractualist methodology,[73] when evaluating whether an agent is permitted to engage in certain conduct, we compare the strength of his reasons for engaging in it, against the reasons that others have to oppose his doing so.[74] In such cases, we typically compare the strength of individuals' opposing reasons against one another.[75]

Such a methodology is especially appropriate for free speech, since it makes sense of the suggestion that some speech is *low-value* in virtue of its disconnection from the values that justify free speech. The notion that there are certain categories of low-value speech is ubiquitous in American constitutional law and among many of its philosophical defenders, explaining why certain categories of speech are afforded weak or no constitutional protection under the First Amendment.[76] But this approach has its critics. As Matthew Kramer notes, "the fatal problem with the high-value/low-value distinction is that the structure of differential levels of protection is profoundly inconsistent with the principle of freedom of expression."[77] On this critique, the whole point of a free speech principle is to forbid governments from making

judgments about what content and especially viewpoints are superior to others, and then granting such speech heightened protection from interference; so how could the application of such a principle depend on such (contentious) judgments?

In reply, I concede that a principle of free speech should dispose us to be seriously prima facie skeptical toward any contention that certain speech is low-value (and thus meriting less protection). We would rightly balk if a court ruled that pro-Muslim or pro-Jewish religious speech was low-value and allowed its suppression on that basis. What is required to show that some speech is low-value is a demonstration that it fails to substantially serve the interests that justify freedom of expression in the first place. When considering a theory of free speech grounded in respect for the development and exercise of certain powers, there is no reason why we cannot inquire into whether there are certain categories of speech that do not serve that underlying rationale.[78]

Overview of Chapters

I have offered an account of the right to free speech, conceived as a weighty natural moral right. Even so, this weighty natural right has limits. This book is dedicated to specifying one important limit on the right, and exploring what its ramifications ought to be for the law and policy of free speech. In so doing, it situates questions of free speech—which have largely been treated as constituting their own special domain of normative inquiry—within broader frameworks in political, moral, and legal philosophy.

Here is a roadmap of the journey ahead. Chapter 2 focuses on the moral responsibilities of speakers. While there are many such responsibilities, I specify a moral duty to refrain from speech that endangers others by promoting clearly wrongful harms, what I call *incitement* for short. The idea that such a duty exists has received surprisingly scant attention in the scholarly literature. Yet it is of foundational importance: the most fundamental question in any debate about restricting speech is whether such speech is wrongful in the first place. This chapter

undertakes that foundational work. On my account, there is a moral duty not to endanger others through speech that encourages clearly wrongful harms. Notably, I argue that incitement can be objectionable even when the speaker does not intend to inspire wrongful harm, even when the incited harm will not occur imminently, and even in cases when the harm isn't likely to eventuate, but would nevertheless be serious if it did. It locates the normative source of the duty to refrain from incitement in a general moral requirement not to harm or endanger others without adequate justification. I suggest that other speech duties (concerning threats, hate speech, misinformation, and self-harm promotion) potentially flow from this underlying source.

Chapter 3 then explores whether incitement, even if prima facie wrongful, is nevertheless protected by the moral right to freedom of speech, as many leading theorists think. I review the range of interests that underpin freedom of expression, which I have just canvassed in this chapter, asking whether they militate in favor of protecting incitement. I argue that many interests are not engaged by incitement. Other interests are engaged, but only weakly. In other cases, interests are substantially engaged but are nevertheless outweighed. The upshot is that the duty to refrain from incitement is not obviated or rendered unenforceable by the right to free speech.

Having thus demonstrated that speakers have an enforceable moral duty to refrain from incitement, I turn in chapter 4 to the obligations of those who *platform* such speech, with a focus on social media platforms. I defend the thesis that platforms have a qualified responsibility to remove incitement. This responsibility is justified by a suite of underlying moral duties: natural rescue duties to defend those endangered by such speech; duties to avoid complicity with users' wrongful speech; and duties to refrain from amplifying innocuous speech that only becomes harmful through amplification.

The book then begins the discussion of whether regulations on harmful speech (again focused on incitement) can be justified all-things-considered. Chapter 5 argues that the justifying purpose of regulations limiting speech is to enforce or specify duties that speakers and platforms owe to others. This duty-based view helps us see why lots of

speech that might be deemed dangerous (such as wartime criticism of foreign policy), or otherwise offensive or destabilizing (such as criticism of politicians), is inherently protected speech. Where speakers and platforms have enforceable duties, they are *potentially liable* to regulation (where to be liable is to lack rights against it). Still, to say that speakers or platforms are potentially liable to regulation enforcing their duties is far from establishing that they *are in fact* liable. That depends on two further criteria: that the regulation does not impose unreasonably excessive costs on its target (the question of *narrow proportionality*) and that there is no better alternative that would achieve the regulation's goals (the question of *necessity*).

If these tests are cleared and liability is established, it still does not follow that the regulation is justified all-things-considered. That depends on a further test, showing that the wider consequences of the regulation on others (usually collateral consequences on bystanders) can be justified (the question of *wide proportionality*). Chapter 5 sets out this overall framework, explores its close connections to similar tests in human rights law, and addresses the narrow proportionality question. The subsequent two chapters tackle necessity and wide proportionality.

Specifically, chapter 6 considers the longstanding view that the right response to harmful speech isn't to ban it; it's to argue back against it. For much controversial and problematic speech, this will be correct; regulation is impermissible on free speech grounds. And even for incontrovertibly harmful speech like incitement, counter-speech is broadly preferable to regulation. The reason, I suggest, is not the commonly stated one: that speech advocating clearly wrongful harm is largely protected speech (a thesis I reject in chapter 3). Rather, the best argument for counter-speech in these cases is supplied by the necessity requirement, which militates against the use of force (or other presumptively wrongful interferences with a right) when a less burdensome but comparably effective option exists. Still, this argument does not rule out the use of regulation. Especially in the online context, we can expect regulation (including private content moderation) to be *more* effective at preventing harms than counter-speech. Moreover, regulation enables

wrongdoers to be held fully accountable for their wrongdoing in ways that ad hoc informal discourse cannot.

The toughest obstacle to justifying regulation, I submit, arises when we consider its collateral consequences (analyzed largely under the criterion of wide proportionality). Chapter 7 catalogues the collateral costs of speech regulations, including the educative benefits of exposure to harmful speech, chilling effects, risks of laws' political misuse and abuse, and worries about inhibiting justified political resistance. While the discussion uses the regulation of incitement as the organizing example, it seeks to illuminate wider issues that arise for the regulation of any category of harmful speech. The tentative and speculative upshot of the chapter is that, in seriously unjust societies, we have strong reason to worry about chilling effects, unjustifiably overbroad restrictions, and the misuse or abuse of government power. Morally minded activists have strong reason to oppose speech regulation in such contexts. However, in robust liberal democracies, my estimation is that collateral costs will not be so large as to render certain speech regulations unjustified. What to do in societies that fall between these two poles will—frustratingly—require some tough judgment calls, about which reasonable people can disagree.

Chapter 8 concludes the book by reflecting on the concrete normative guidance it offers to a range of actors: from legislators to social media platforms to ordinary citizens. The moral vision that emerges from the book is a complex, multitactical, all-hands-on-deck effort to defuse the dangers of harmful speech while upholding freedom of expression, properly understood. The task of combatting speech that incites harm is one that falls on us all.

INDEX