Canadian Journal of Law & Jurisprudence XXXVIII No.1 February 2025, 29-67

 doi: 10.1017/cjlj.2024.20

The Free Speech Principle: A Topography

Alexander Brown
University of East Anglia (UEA), Norwich, UK
Email: alexander.c.brown@uea.ac.uk

Abstract

This paper argues that what scholars call 'the free speech principle' is not one principle but a slew of principles, and that these principles harbour several important differences that have remained largely unremarked upon, namely: (i) extending vs. limiting principles; (ii) comparative vs. non-comparative principles; and (iii) monistic vs. pluralistic principles. The paper also critically assesses certain generalisations that people might be tempted to make about these different principles, such as that one kind of free speech principle is harder to defend than another. Finally, the paper teases out the practical as well as theoretical implications of these insights, including degrees of complexity, the logical relationship between free speech principles and free speech policy dilemmas, and the virtue of compromise over free speech principles.

Keywords: *Free speech; Normative Principles; Compromise; Hate Speech*

I. Introduction

It is striking how, on the one hand, uses of the phrase 'the free speech principle' and similar locutions[1] are ubiquitous within the philosophical literature on free speech but, on the other hand, no two articulations of the free speech principle are the same, and many do not even fall in the same ballpark as one another.[2] It is also noteworthy how often philosophers speak of 'the free speech principle' without mentioning, much less critically examining, key differences between the many versions of the principle found in the literature,[3] although there are some

1. Such as 'the principle of free speech', 'the principle of freedom of expression', and 'the shibboleth of free speech'.
2. See e.g. Thomas Scanlon, "A Theory of Freedom of Expression" (1972) 1:2 Philosophy & Public Affairs 204 at 209, 213; Frederick Schauer, *Free Speech: A Philosophical Enquiry* (Cambridge University Press, 1982) at 7-8; Wojciech Sadurski, *Freedom of Speech and Its Limits* (Kluwer Academic, 1999) at 2, 4; Larry Alexander, *Is There a Right of Freedom of Expression?* (Cambridge University Press, 2005) at 38; Steven J Heyman, *Free Speech and Human Dignity* (Yale University Press, 2008) at 44-45; Nigel Warburton, *Free Speech: A Very Short Introduction* (Oxford University Press, 2009) at 1-2; Jeffrey W Howard, "Dangerous Speech" (2019) 47:2 Philosophy & Public Affairs 208 at 210; Matthew H Kramer, *Freedom of Expression as Self-Restraint* (Oxford University Press, 2021) at 1.
3. See generally *supra* note 2.

exceptions to this tendency.[4] Now I do not mean to suggest that scholars are unaware of and do not write about their disagreements over free speech, especially their disagreements about why free speech matters, its potential limits, and relevant free speech dilemmas, obligations, and policies. Clearly, scholars routinely both engage in and reflect on such disagreements. Rather, I am interested in the free speech principle as a particular type of propositional format or technical device for framing free speech issues.

This article argues that the free speech principle is not one principle but a slew of principles, and that these principles harbour several important differences that have remained largely unremarked upon. To clear up conflations and confusions, this article provides a conceptual map or topography of the different kinds of free speech principles circulating in the academic literature. In particular, I introduce and analyse three distinctions: (i) extending vs. limiting principles; (ii) comparative vs. non-comparative principles; and (iii) monistic vs. pluralistic principles. These distinctions underscore the heterogeneity of free speech principles—principles that operate at different levels of generality and abstraction and have different philosophical cruxes.

Furthermore, I address certain initially plausible beliefs, assumptions, and generalisations that people might have, explicitly or implicitly, about these different kinds of free speech principles. These generalisations (*qua* meta-level principles) may be either actual or hypothetical for the purposes of what is at heart a philosophical exercise. Recall John Stuart Mill's observation from *On Liberty*:

> So essential is this discipline [of critical thought] to a real understanding of moral and human subjects, that if opponents of all important truths do not exist, it is indispensable to imagine them and supply them with the strongest arguments which the most skilful devil's advocate can conjure up.[5]

More specifically, I critically assess generalisations of the form that one kind of free speech principle is more demanding, more ambitious, and ultimately harder to defend than another kind; and conversely that some kinds of free speech principles are an easier option. I argue that some such generalisations are both true and useful, but not all of them are.

My argument is socially, politically, and legally important because of the growing prominence given over to the free speech principle in public life.

---

4. There are scholars who at least sound a note of caution or express general scepticism about the phrase 'the free speech principle' and/or always take care to speak of different free speech principles. See Kent Greenawalt, *Speech, Crime, and the Uses of Language* (Oxford University Press, 1989) at 9; David A Strauss, "Persuasion, Autonomy, and Freedom of Expression" (1991) 91:2 Colum L Rev 334 at 357, n 64; Steven Shiffrin, "The Politics of the Mass Media and the Free Speech Principle" (1994) 69:3 Ind LJ 689 at 692, n 20; Stanley Fish, *There's No Such Thing As Free Speech … and it's a good thing, too* (Oxford University Press, 1994) at 102; Robert Post, "Recuperating First Amendment Doctrine" (1995) 47:6 Stan L Rev 1249 at 1272; Susan J Brison, "The Autonomy Defense of Free Speech" (1998) 108:2 Ethics 312 at 320.

5. John Stuart Mill, "On Liberty" in *Utilitarianism, On Liberty, Considerations on Representative Government*, ed by Geraint Williams (JM Dent Orion, 1993) at 105.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

This prominence can be seen in everything from this principle becoming a marker of personal or even national or patriotic identity in some political cultures;[6] through to the role played by the free speech principle as a political football between parties and movements and as a hot button issue in wider culture wars; to the free speech principle being one of the main philosophical ideas influencing the development of law in a growing number of areas. Understanding that (i) the free speech principle is not a monolith but a cluster of principles, (ii) that there are important differences between these principles, and (iii) that some principles are harder to defend than others, but also that some principles only appear harder to defend than others, is key to clearing up muddles and incorrect assumptions that can hamper or skew these social, political, and legal debates. To help draw out some of the differences between free speech principles and the increasing importance of the idea of the free speech principle in public life—as well as to give the article a focus and golden thread—I shall concentrate on a single unifying dilemma concerning free speech and its limits, namely, hate speech laws. Few free speech issues today are as hotly contested and as routinely misunderstood.

The remainder of the article is structured as follows. In Section II, I explain what is meant by 'the free speech principle' in the academic literature, including demarcating philosophical from doctrinal principles.

Section III introduces and analyses the distinction between free speech principles that are extending (support the protection of free speech) and those that are limiting (justify limits on free speech). I argue that, based on conventions in the relevant literature about differential burdens of argument for supporting free speech protections and justifying limitations on speech respectively, it can seem much easier to defend extending principles than limiting principles. Focusing on the example of hate speech laws, I aim to debunk this conventional wisdom by showing that differential burdens of argument are based on arbitrary double standards (an absence of good reasons). It might be possible to come up with good reasons, but doing so is not the easy option by any means.

Section IV looks at the distinction between comparative and non-comparative free speech principles. In this case, I critically assess the generalisation that it is harder to defend non-comparative than comparative principles. I argue this generalisation does not hold water—for example, coming up with criteria for comparing the harms of speech and non-speech is no easier than coming up with theories of harm thresholds in relation to speech.

Section V delves into the distinction between monistic and pluralistic free speech principles. Here I argue that it *is* possible to stand up the generalisation that it is harder to defend monistic than pluralistic principles. One reason for this is that monistic principles typically contain the quantifiers 'the only' or 'the most important' and so need to prove that one thing and one thing alone matters or else that one thing matters more than any other. I also canvass and reject one seemingly ingenious solution to defending harm-based monistic free speech

6. See Steven H Shiffrin, *Dissent, Injustice, and the Meanings of America* (Princeton University Press, 1999).

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

principles, namely, to embrace a dependent theory of harm according to which 'harm' is simply the label we give to something for which there is a pro tanto reason to restrict the speech that constitutes or produces it. I argue this approach is no silver bullet for defending monistic principles, due to several major flaws in the dependent account.

Finally, Section VI sets forth what I take to be the important practical as well as theorical implications of uncovering the truth that the free speech principle is not a monolith but instead a cluster of philosophical principles. These implications have to do with degrees of complexity, the logical relationship between free speech principles and free speech policy dilemmas, and the virtue of compromise over free speech principles. To briefly expand on the last of these implications: If there is a plethora of free speech principles and if there is reasonable disagreement over them (such that there is no sovereign free speech principle upon which people can agree), then this highlights the need for compromise over free speech principles, and in turn the importance of further research into the philosophy of compromise.

II. What is a Free Speech Principle?

When I use the phrase 'the free speech principle', I do not mean to use a shorthand label for a particular constitutional guarantee and/or human right to freedom of expression, such as can be found in written texts or documents of domestic and international law. Nor do I mean a shorthand label for certain doctrinal free speech principles that can be found in the case law and/or explanatory memorandums pertaining to constitutional and/or human rights to freedom of expression. The doctrinal free speech principle (or set of principles) connected with the First Amendment to the United States Constitution has been extensively explored and critically assessed by American legal scholars, for example.[7]

Instead, I mean to refer to a philosophical principle (or set of principles) typically articulated by legal theorists and legal and political philosophers alike, and which is a postulate of normative jurisprudence.[8] Free speech principles in this sense are supposed to be general guides to a range of public policy decisions (legislative, judicial, constitutional) about the protection or restriction of speech based on standards of right and good that have broad social, political, ethical, and

7. See e.g. Alexander Meiklejohn, *Free Speech And Its Relation to Self-Government* (Harper & Brothers, 1948); Thomas I Emerson, *The System of Freedom of Expression* (Random House, 1970); C Edwin Baker, *Human Liberty and Freedom of Speech* (Oxford University Press, 1989); James Weinstein, *Hate Speech, Pornography, and the Radical Attack on Free Speech Doctrine* (Westview Press, 1999); Steven H Shiffrin, *What's Wrong with the First Amendment?* (Cambridge University Press, 2016).

8. For the distinction between philosophical and doctrinal free speech principles, see also Greenawalt, *supra* note 4 at 3; Frederick Schauer, "What Is Speech? The Question of Coverage" in Adrienne Stone & Frederick Schauer, eds, *The Oxford Handbook of Freedom of Speech* (Oxford University Press, 2021) 158 at 165.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

moral dimensions, as opposed to shorthand descriptions of what some particular body of law, court, or entire system of law says and does about free speech.[9]

Of course, the phrase 'the free speech principle' and similar locutions is used not merely by legal theorists and legal and political philosophers, but also by politicians, civil society groups, human rights organisations, social media platforms, journalists, writers, and the public at large. But the focus in this article is on uses of the phrase 'the free speech principle' within the academic canon as opposed to public discourse in general.

As I understand it, the phrase 'the free speech principle' suggests a particular type of propositional format or technical device for framing free speech issues. For the sake of clarity, I propose that this format or device should be thought of as possessing five identifying characteristics. First, the phrase 'the free speech principle' suggests a formal and canonical statement about free speech in the sense of something that is veridical, reasonably brief, and purports to be not merely accurate but also something to which relevant agents can and should turn as an authority on the subject. In other words, free speech principles are typically succinct statements about free speech that are intended to serve as fundamental truths and the bases for a system of wider beliefs and chains of reasoning about the subject.

Second, free speech principles normally have a high level of abstraction, meaning they tend more towards the general than the particular. In other words, they usually take the form of rules of thumb about free speech that cover a range of circumstances and cases, or even all circumstances and cases.

Third, free speech principles typically contain declaratory, justificatory, and/or prescriptive content about free speech (at least two out of the three). For some free speech principles, for example, this means not merely do they declare that speech should be protected or should not be restricted, but they also provide a deeper understanding of when and why this holds true and/or what this means in terms of which decisions should be taken and what policies ought to be pursued.

Fourth, free speech principles tend to be normative, in the sense that they are intended to establish a norm or standard that relevant agents have a reason to adhere to because it is a principle. Typically, this also means relevant agents can expect other agents to adhere to the principle and to be responsive to that fact. A free speech principle in this sense can be accepted by relevant agents (such as governments, institutions, organisations, officials, judges) as a reason for taking a particular decision or adopting a certain policy on free speech (e.g., 'We decided not to ban this speech out of respect for the free speech principle') but can also serve as a benchmark for evaluating an agent's decision or policy

9. Of course, according to some philosophies of law, what a particular body of law, court, or entire system of law says and does is itself rooted in broader ethical and moral principles. See Ronald Dworkin, *Freedom's Law: The Moral Reading of the American Constitution* (Harvard University Press, 1996) [Dworkin, *Freedom's Law*]; Ronald Dworkin, *Justice for Hedgehogs* (Harvard University Press, 2011).

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

on free speech (e.g., 'They ought not to have banned this speech, because they should have respected the free speech principle').

The fifth desideratum of the free speech principle (or set of principles) is that it should be capable of standing as a special principle, in the sense that it has a distinctive quality that flows from and reflects the special quality of free speech itself. This is perhaps the most controversial desideratum on the list, because some scholars (free speech sceptics) deny that free speech has or ought to have a high degree of distinctiveness. I shall return to this scepticism throughout the article, but for now I leave this desideratum in place, albeit with a metaphorical asterisk placed next to it.

There is one further clarification I need to make before beginning in earnest. Although this article is intended to provide a reasonably comprehensive conceptual map of free speech principles, I shall not attempt to discuss every kind of free speech principle imaginable. In particular, I shall concentrate on object-level free speech principles—namely, principles that take free speech as their direct subject matter. This includes higher level free speech principles that say something about which values support free speech, or that specify what sort of purposes can justify rightful limitations on speech. By contrast, meta-level free speech principles make clarifications, qualifications, and generalisations about object-level free speech principles (i.e., principles about principles). For example, a meta-level free speech principle might consist of a "principle for defining the scope of freedom of expression."[10] Such a principle could clarify which phenomena should or should not count—for the purpose of understanding and applying object-level free speech principles—as speech, restrictions on speech, agents culpable for restrictions on speech, agents capable of speech in the relevant sense, and social practices or contexts in which restrictions of speech matter. Important differences between such principles have been covered elsewhere in the literature and so I shall not evaluate them here.[11] However, there are some meta-level free speech

10. Alexander, *supra* note 2 at 38.

11. One key distinction among principles for defining the scope of freedom of expression is between principles that involve 'defining-in' and principles that involve 'defining-out' (or 'adding' and 'subtracting'). This can be the difference between, on the one hand, a principle (meta-level) that starts from the position that nothing should be counted as 'speech' within a free speech principle (object-level) unless and until there is good reason to define-in or add things and, on the other hand, a principle (meta-level) that starts from the opposite position that (almost) everything should be counted as 'speech' within a free speech principle (object-level) unless and until there is good reason to define-out or subtract things. See Frederick Schauer, "Categories and the First Amendment: A Play in Three Acts" (1981) 34:2 Vand L Rev 265; Robert Mark Simpson, "Defining 'Speech': Subtraction, Addition, and Division" (2016) 29:2 Can JL & Jur 457.

Another important distinction among principles for defining the scope of freedom of expression is between dependent and independent principles. This can be the difference between, on the one hand, a principle (meta-level) that says valid reasons for classifying or not classifying certain things as 'speech' within a free speech principle (object-level) are dependent on prior reasons for protecting or restricting the things in question and, on the other hand, a principle (meta-level) that says valid reasons for classifying or not classifying certain things as 'speech' within a free speech principle (object-level) can include an independent or freestanding conception of what speech is. See Schauer, *supra* note 2 at 89-92; Greenawalt, *supra* note 4 at 12, 40; Lawrence Alexander & Paul Horton, "The Impossibility of a Free Speech Principle Review

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

principles I shall critically assess. As stated above, I shall critically assess generalisations (actual or hypothetical) to the effect that a certain kind of free speech principle is harder to defend than another.

III. Extending vs. Limiting Principles

The first distinction I want to analyse is between free speech principles that extend and those that limit free speech. Starting with the former, I use the term 'extend' in a deliberately broad way to mean principles that justify free speech as well as those that specify the range of protections that should be given to free speech, and those that defend states of affairs in which free speech occupies or covers a wider area. This is a broad category that includes several subtypes of free speech principles.

One subtype of extending principles sets forth certain rationales, purposes, or justifications for protecting speech such as those based on interests, rights, duties, public goods, or fundamental values. The academic literature on free speech

---

Essay" (1983) 78:5 Nw UL Rev 1319 at 1322; David Braddon-Mitchell & Caroline West, "What is Free Speech?" (2004) 12:4 J Political Philosophy 437 at 439; Schauer, *supra* note 8 at 165.

These different kinds of meta-level principles can be seen at play, explicitly or implicitly, in academic debates about hate speech and in particular about whether the term 'speech' in 'free speech' should cover the referents of the term 'speech' in 'hate speech'. For example, some scholars argue that the use of racial slurs and other recognisable forms of hate speech on college campuses and in other social situations should sometimes not be counted as 'speech' in the technical sense of the word for the purposes of applying free speech principles. They argue the reasons for protecting free speech scarcely apply to some hate speech, and also that it is much harder to justify campus codes, for example, if the problem they are supposed to tackle is classified as 'transgressive speech'. They argue it is in keeping with the justifications for hate speech regulations to classify at least some hate speech as 'non-speech', that is, as 'transgressive behaviour', including 'fighting words', 'threats', 'situation-altering conduct', 'discriminatory harassment', 'acts of silencing', 'subordination', and 'incitement to hatred, discrimination, and violence'. See Kent Greenawalt, *Fighting Words: Individuals, Communities, and Liberties of Speech* (Princeton University Press, 1995); Kenneth L Marcus, "Higher Education, Harassment, and First Amendment Opportunism" (2008) 16:4 William & Mary Bill of Rights J 1025; Alexander Tsesis, "Campus Speech and Harassment" (2017) 101:5 Minn L Rev 1863; Ishani Maitra & Mary Kate McGowan, "On Racist Hate Speech and the Scope of a Free Speech Principle" (2010) 23:2 Can JL & Jur 343.

By contrast, other scholars are more open to seeing the use of racist slurs and other hate speech on college campuses as still 'speech' in a sense relevant to applying principles of free speech. This does not mean to say they conclude that campus codes and other hate speech regulations can never be justified whilst respecting free speech principles. Indeed, they argue there are weighty reasons in favour of regulating the sort of hate speech that transgresses civility norms, norms of public discourse, and ideals of deliberative democracy, for example. Rather, it means that these scholars see instances of hate speech as genuine hard cases in which there is a need to somehow balance or otherwise reconcile reasons in favour of regulating certain kinds of 'speech' with respect for principles of free speech. See Cass R Sunstein, *Democracy and the Problem of Free Speech* (Free Press, 1993) at ch 6, ch 8; Suzanne Rice, "'Hate Speech' and the Need for Moral Standards in Communicative Interaction" (1994) 50 Philosophy Education 70; Patricia S Mann, "Hate Speech, Freedom, and Discourse Ethics in the Academy" in David S Caudill & Steven Jay Gold, eds, *Radical Philosophy of Law: Contemporary Challenges to Mainstream Legal Theory and Practice* (Humanities Press, 1995) 255; James E Fleming, "Securing Deliberative Democracy"(2004) 72:5 Fordham L Rev 1435.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

contains principles invoking a wide array of different justifications,[12] including but not limited to the following: access to information;[13] truth discovery and knowledge acquisition (epistemic goods);[14] autonomy;[15] self-realisation;[16] the free development and operation of the thinking mind;[17] ideals of communicative action;[18] human dignity;[19] equal concern and respect;[20] the value of dissent;[21] the checking value;[22] ideals of democracy (representative, participatory, deliberative);[23] and political legitimacy.[24]

A second subtype of extending principles sets forth how certain rationales, purposes, or justifications for free speech relate to the relevant protections. For example, a free speech principle might say there exists at least one interest, right, duty, public good, or value that uniquely justifies free speech (meaning it does not justify other sorts of freedoms in the same way), such that free speech deserves special protection over and above the protections afforded to other freedoms. Even if free speech is justified by roughly the same stock of morally relevant features as justify some or all other freedoms, a free speech principle might assert that a certain feature has a "special relationship"[25] to speech or applies

12. See Schauer, *supra* note 2; Alan Haworth, *Free Speech: All That Matters* (John Murray Press, 2015); Matteo Bonotti & Jonathan Seglow, *Free Speech* (John Wiley, 2021).
13. See Daniel A Farber, "Free Speech without Romance: Public Choice and the First Amendment" (1991) 105:2 Harv L Rev 554.
14. See Mill, *supra* note 5; Eugene Volokh, "In Defense of the Marketplace of Ideas/Search for Truth as a Theory of Free Speech Protection" (2011) 97:3 Va L Rev 595.
15. See Baker, *supra* note 7; C Edwin Baker, "Autonomy and Free Speech" (2011) 27:2 Const Commentary 251; Brison, *supra* note 4.
16. See Thomas I Emerson, "Toward a General Theory of the First Amendment" (1963) 72:5 Yale LJ 877; Martin H Redish, "The Value of Free Speech" (1982) 130:3 U Pa L Rev 591.
17. See Seana Valentine Shiffrin, "A Thinker-Based Approach to Freedom of Speech" (2011) 27: Const Commentary 283 at 287.
18. See Lawrence Byard Solum, "Freedom of Communicative Action: A Theory of the First Amendment Freedom of Speech" (1989) 83:1 Nw UL Rev 54
19. See Heyman, *supra* note 2.
20. See Dworkin, *Freedom's Law, supra* note 9.
21. See Shiffrin, *supra* note 6.
22. See Vincent Blasi, "The Checking Value in First Amendment Theory" (1977) 2:3 American Bar Foundation Research J 521.
23. See Alexander Meiklejohn, *Political Freedom: The Constitutional Powers of the People* (Harper & Brothers, 1960); Alexander Meiklejohn, "The First Amendment Is An Absolute" (1961) 1961 Supreme Court Rev 245; Sunstein, *supra* note 11; Robert Post, "Democracy and Equality" (2005) 1:2 Law, Culture & Humanities 142; Robert Post "Participatory Democracy and Free Speech" (2011) 97:3 Va L Rev 477; James Weinstein, "Participatory Democracy as the Central Value of American Free Speech Doctrine" (2011) 97:3 Va L Rev 491; Eric Heinze, *Hate Speech and Democratic Citizenship* (Oxford University Press, 2016); Maxime Lepoutre, *Democratic Speech in Divided Times* (Oxford University Press, 2021).
24. See C Edwin Baker, "Autonomy and Hate Speech" in Ivan Hare & James Weinstein, eds, *Extreme Speech and Democracy* (Oxford University Press, 2009) 139; Ronald Dworkin, "Foreword" in Hare & Weinstein, *supra* note 24, v ; Ronald Dworkin, "Reply to Jeremy Waldron" in Michael Herz & Peter Molnar, eds, *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge University Press, 2012) 342; James Weinstein, "Hate Speech Bans, Democracy, and Political Legitimacy" (2017) 32:3 Const Commentary 527.
25. Leslie Kendrick "Free Speech as a Special Right" (2017) 45:2 Philosophy & Public Affairs 87 at 92.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

"with special force" to speech,[26] such that free speech merits being considered a special right. For example:

> Perhaps free speech and intimate association [sexual freedom] have equally essential but differing relationships to autonomy. . . . These activities are distinctive enough to be identified separately, and thus to count as special rights.[27]

In addition, a free speech principle might say something about the sort of protection free speech deserves. For example:

> Protected speech appears as a distinct ideal from unhindered speech as soon as we recognize that the point of protection is not to make interference less attractive and less probable but to interfere with the very possibility of interference: to remove that option altogether or to replace it by a penalized alternative. Protecting you means erecting obstacles to the interference of other people in any scenarios, however improbable, where they might choose to try to interfere. And that is quite distinct from trying to make their interference less probable.[28]

At an even more fine-grained level, a free speech principle might specify the particular protections that free speech merits, such as certain constraints on judicial discretion. For example:

> [T]aking stock of the legal system's own limitations, we must realize that judges, being human, will not only make mistakes but will sometimes succumb to the pressures exerted by the government to allow restraints that ought not to be allowed. To guard against these possibilities we must give judges as little room to maneuver as possible and, again, extend the boundary of the realm of protected speech into the hinterlands of speech in order to minimize the potential harm from judicial miscalculation and misdeeds.[29]

A third subtype of extending principles are negative principles that specify the sorts of rationales, purposes, or justifications for restricting speech that are *not* acceptable, adequate, legitimate, sufficiently strong, or warranted by the lights of the principle. The legal philosopher Matthew Kramer is describing this sort of principle when he writes: "the principle of freedom of expression imposes absolute restrictions on the purposes—the ends and means—that can legitimately be pursued by a system of governance through any measures that prohibit types or instances of communicative conduct."[30] Numerous examples of this subtype of extending principles can be found in the academic literature (as well as in relevant bodies of law), and together they specify a wide range of justifications for restricting speech that are *not* legitimate, including but not limited to the following:

---

26. Kent Greenawalt, "Free Speech Justifications" (1989) 89:1 Colum L Rev 119 at 120.
27. Kendrick, *supra* note 25 at 102 [footnote omitted].
28. Philip Pettit, "Two Concepts of Free Speech" in Jennifer Lackey, ed, *Academic Freedom* (Oxford University Press, 2018) 61 at 64.
29. Lee C Bollinger, *The Tolerant Society: Freedom of Speech and Extremist Speech in America* (Oxford University Press, 1986) at 78.
30. Kramer, *supra* note 2 at 1.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

"[preserving] what shall be orthodox in politics, nationalism, religion, or other matters of opinion";[31] "simply because society finds the idea itself offensive or disagreeable";[32] "just because these tastes or opinions disgust those who have the power to shut him up or lock him up";[33] "'information' or 'ideas' that ... offend, shock or disturb the State or any sector of the population";[34] "simply because it is upsetting or arouses contempt";[35] "that their officials or fellow-citizens think that their opinions about the right way for them to lead their own lives are ignoble or wrong";[36] "the hypothesis that the attitudes ... are demeaning or bestial or otherwise unsuitable to human beings of the best sort";[37] "the communicative tenor of the conduct";[38] the enforcement of "civility norms";[39] "harms to certain individuals which consist in their coming to have false beliefs as a result of those acts of expression ... [and] harmful consequences of acts performed as a result of those acts of expression, where ... the act of expression led the agents to believe (or increased their tendency to believe) these acts to be worth performing";[40] "because people would be persuaded by it";[41] "that the speech is likely to persuade people to do something that the government considers harmful";[42] "opinions that the government finds wrong or even hateful";[43] "the speech expresses hateful thoughts toward certain individuals and groups ... [and] all of us who believe that all human beings are entitled to full and equal rights hate the diametrically different thoughts expressed by [such speech]";[44] and "hateful viewpoints [that] ... directly challenge value democracy's commitment to freedom and equality."[45]

Turning next to limiting principles, these are principles that, although they start with the assumption that speech deserves protection, at the same time specify the limits of free speech. Once again, there are different subtypes of limiting principles. One subtype is principles that say something in general about the limited nature of protection that should be afforded to free speech and which, therefore, open the door to permissible limits or restrictions on expression. For example, a free speech principle might state that the right to free speech is derived from a more fundamental right protecting freedom in general; that this general

31. *West Virginia Board of Education v Barnette*, 319 US 624 at 642 (1943).
32. *Texas v Johnson*, 491 US 397 at 414 (1989).
33. Dworkin, *Freedom's Law*, *supra* note 9 at 238.
34. *Handyside v United Kingdom* (1976) ECHR (Ser A4) 5493/72, 1 EHRR 737 at para 49.
35. *Snyder v Phelps*, 562 US 443 at 458 (2011).
36. Ronald Dworkin, "Is There a Right to Pornography?" (1981) 1:2 Oxford J Leg Stud 177at 194.
37. *Ibid* at 195.
38. Kramer, *supra* note 2 at 35.
39. Robert Post, "Hate Speech" in Hare & Weinstein, *supra* note 24, 123 at 136.
40. Scanlon, *supra* note 2 at 213.
41. Sunstein, *supra* note 11 at 162.
42. Strauss, *supra* note 4 at 335.
43. *American Booksellers Association v Hudnut*, 771 F (2d) 323 at 328 (7th Cir 1985).
44. Nadine Strossen, "Incitement to Hatred: Should There Be a Limit?" (2001) 25 S Ill ULJ 243 at 244.
45. Corey Brettschneider, "Value Democracy as the Basis for Viewpoint Neutrality: A Theory of Free Speech and Its Implications for the State Speech and Limited Public Forum Doctrines" (2013) 107:2 Nw UL Rev 603 at 611.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

right to freedom is a non-absolute right, meaning freedom may be limited or restricted; and that, therefore, the right to free speech is likewise subject to limitation or restriction. Consider the following minimalist free speech principle: There should exist a substantial sphere of personal liberty within which the individual ought to be free from governmental interference, but this sphere is limited in nature and does not cover everything an individual does or says. Insofar as certain speech falls within this sphere of personal liberty, then it follows that an individual's speech should be free from governmental interference, but insofar as certain speech falls outside of this sphere of personal liberty, then an individual's speech may be limited or restricted.

A second subtype of limiting principles is exemplified by principles setting out specific rationales, purposes, or justifications for restricting speech which the principles declare to be acceptable, adequate, legitimate, sufficiently strong, or warranted. Once again, many examples of this subtype of limiting principles can be found in the academic literature (as well as in relevant bodies) of law, but to make the illustrations manageable and to give them a focus, I shall only mention justifications for restricting hate speech. Even here, relevant justifications cover a wide spectrum, including but not limited to the following: psychological harm;[46] personal security;[47] oppression and subordination;[48] personal development;[49] 'necessary in a democratic society' (interests of public safety, the prevention of disorder or crime, and the protection of the reputation of others);[50] human dignity;[51] autonomy;[52] the public goods of cultural diversity and intercultural

46. See Mari J Matsuda, "Public Response to Racist Speech: Considering the Victim's Story" (1989) 87:8 Mich L Rev 2320; Richard Delgado & Jean Stefancic, *Understanding Words That Wound* (Routledge, 2004); Alexander Brown, *Hate Speech Law: A Philosophical Examination* (Routledge, 2015) at ch 3.1.

47. See Brown, *supra* note 46 at ch 3.3; Alexander Tsesis, *Destructive Messages: How Hate Speech Paves the Way for Harmful Social Movements* (New York University Press, 2002).

48. See Mary Kate McGowan, "On 'Whites Only' Signs and Racist Hate Speech: Verbal Acts of Racial Discrimination" in Ishani Maitra & Mary Kate McGowan, eds, *Speech and Harm: Controversies Over Free Speech* (Oxford University Press, 2012) 121; Rae Langton, "Beyond Belief: Pragmatics in Hate Speech and Pornography" in Maitra & McGowan, *supra* note 48, 72; Brown, *supra* note 46 at ch 3.4, 3.5; Ishani Maitra & Mary Kate McGowan, "Language and Free Speech" in Justin Khoo & Rachel Katharine Sterken, eds, *The Routledge Handbook of Social and Political Philosophy of Language* (Routledge, 2021) 317; Katharine Gelber, "Differentiating Hate Speech: A Systemic Discrimination Approach" (2021) 24:4 Critical Rev Intl Social & Political Philosophy 393.

49. See Brown, *supra* note 46 at ch 4.

50. See *Glimmerveen and Hagenbeek v The Netherlands* (1979), 8348/78 ECHR 8; *Jersild v Denmark* (1994), 15890/89 ECHR 33, 19 EHRR 1 ; *Garaudy v France* (dec), No 65831/01 (7 July 2003); *Soulas and others v France*, No 15948/03 (10 October 2008); *Balsyte-Lideikiene v Lithuania*, No 72596/01 [2008] ECHR 1195; *Féret v Belgium*, No 15615/07 (16 July, 2009); *Willem v France*, No 10883/05 (16 July 2009); *Aksu v Turkey*, No 4149/04, [2012] ECHR 445, 56 EHRR 4; *Vejdeland and Others v Sweden*, No 1813/07, [2012] ECHR 242, 58 EHRR 15; *Atamanchuk v Russia*, No 4493/11, [2020] ECHR 133; *Sanchez v France*, No 45581/15, [2021] ECHR 724.

51. See Heyman, *supra* note 2 at ch 10; Brown, *supra* note 46 at ch 3.6.

52. See Brison, *supra* note 4; Brown, *supra* note 46 at ch 3.2.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

dialogue;[53] the assurance of civic dignity for all;[54] full and equal opportunity to participate in public debate;[55] real access to the formation of public opinion for all;[56] and political legitimacy.[57]

My proposed analysis of the differences between extending and limiting free speech principles is comprised of three main points. My first point is that extending principles and limiting principles can be—but are not necessarily—two sides of the same coin. For one thing, extending principles do not always explicitly and directly articulate (and might not even imply) where support for protecting free speech ends and where possible justifications for restrictions might begin. That being said, it is also obvious that free speech scholars can, and often do, develop and articulate extending principles and limiting principles in conjunction, such as under the umbrella of a 'theory' of free speech. For example, a theory might justify free speech based on the value of autonomy but at the same time specify when restrictions on speech are legitimate for the sake of protecting autonomy.[58] This practice does not, however, make the distinction between the principles redundant. For these principles are logically distinct entities. Thus, a scholar who defends an eclectic theory of free speech could coherently combine extending principles which appeal to one value (or set of values) with limiting principles that appeal to another value (or set of values). This might be done in relation to a theory of free speech and its limits relating to hate speech, for example.[59]

My second point (which is also connected to the first) relates to the relationship between free speech principles and free speech absolutism. Free speech absolutism is, according to one possible reading, the view that the right to free speech is absolute and unqualified in the twin senses that it should never be put on the balancing scales with, or made to defer to, whatever policy goals, pressing social needs, or political ends a public authority may wish to pursue, and is not qualified by other rights, such that in effect a public authority can never justify breaching the right to free speech.[60] The point I wish to make here is that extending principles could in theory be used to articulate or defend either absolutist or non-absolutist visions of free speech, whereas, commonsensically, limiting principles can only ever be associated with non-absolutist visions of free speech.[61] This is because limiting principles declare either in general terms that

53. See Brown, *supra* note 46 at ch 6.
54. See Jeremy Waldron, "2009 Oliver Wendell Holmes Lectures: Dignity and Defamation: The Visibility of Hate" (2010) 123:7 Harv L Rev 1597; Brown, *supra* note 46 at ch 5.
55. See Owen Fiss, *The Irony of Free Speech* (Harvard University Press, 1996).
56. See Brown, *supra* note 46 at 194-201.
57. See *ibid* at ch 7.2; Alexander Brown, "Hate Speech Laws, Legitimacy, and Precaution: Reply to Weinstein" (2017) 32:3 Const Commentary 599.
58. See e.g. Brison, *supra* note 4.
59. See Tsesis, *supra* note 47; Brown, *supra* note 46.
60. For a distinction between different senses of absolutism, see Frederick F Schauer, "Must Speech Be Special?" (1983) 78:5 Nw UL Rev 1284 at 1285, n 9.
61. For an illustration of extending principles being put into service of an absolutist vision of free speech, see Baker, *supra* note 7; Volokh, *supra* note 14. For an illustration of extending principles being used to support a non-absolutist vision of free speech, see Weinstein, *supra* note 23; Weinstein, *supra* note 24.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

limitations on speech can be legitimate, or else provide specifics about when and why certain limitations on speech can be legitimate; but these are limitations no free speech absolutist could countenance.

My third point is that there exists a perception within some of the academic literature on free speech, although other scholars take a different view,[62] that the burden of argument falls more heavily on those seeking to justify limitations on free speech than those seeking to defend the protection of free speech in the first place.[63] This is the idea that there is a presumption in favour of free speech: the convention that it is somehow axiomatic that speech deserves protection, whereas limitations on speech require special justification.[64] As a result, it can appear easier to defend extending principles than limiting principles. By way of illustration, on the one hand, it can be relatively straightforward to come up with certain values that would be well-served by speech protections, especially if one does not necessarily have to demonstrate that those values require free speech above all else, and if one is not obliged to show that those values cannot be served by other things besides free speech. On the other hand, when it comes to demonstrating that certain values justify limitations on speech, the conventional burden of argument involves showing that those values necessitate limitations on speech and that the relevant limitations on speech are the least restrictive means of serving those values. Furthermore, the burden of argument in the case of some extending principles demands only that someone can come up with arguments about why certain values would be well-served by free speech without having to consider the ways in which other values might suffer as a result of unfettered free speech. Conversely, the burden of argument in the case of some limiting principles seems to be higher, demanding that someone must come up with not only arguments about why certain values necessitate limitations on speech but also arguments about why these values are more important or weightier than those favouring the protection of free speech.

Some of these conventions concerning differential burdens of argument have been relied upon, explicitly or implicitly, by scholars who argue against hate speech laws, for example. They have claimed that hate speech laws ought to satisfy but cannot satisfy special constraints on limiting free speech, such as relating to the least restrictive alternative, unintended consequences, and slippery slopes, thereby placing a higher burden of argument on defenders of hate speech laws.[65] However, the conventions themselves are rarely, if ever, defended with deeper philosophical justifications, i.e., justifications that go all the way down to the fundamental level of justifying why there ought to be differential burdens of

62. See e.g. Alexander, *supra* note 2.

63. See Gehan Gunatilleke, "Justifying Limitations on the Freedom of Expression" (2021) 22:1 Human Rights Rev 91 at 93.

64. Perhaps a more subtle version of this convention has it that there is a low burden of argument to justify special protection for free speech given the baseline of a presumption of liberty, whilst there is a high burden of argument to justify limiting free speech given the baseline of special protection for free speech.

65. For citations to this literature, see Brown, *supra* note 46 at ch 9; Alexander Brown & Adriana Sinclair, *The Politics of Hate Speech Laws* (Routledge, 2020) at chs 5-6.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

argument in the first place. Yet in the absence of such justifications, differential burdens of argument can be seen as arbitrary double standards. What sound reasons are there to impose different adequacy constraints on extending and limiting principles respectively, such that extending principles can be adequate even when effectively imposing low burdens of argument on those supporting free speech protections, whereas, in order to be adequate, limiting principles must impose high burdens of argument when it comes to justifying limitations on speech? Now, I am not suggesting that it would be impossible to furnish sound reasons in defence of these differential burdens of argument. Rather, I am saying that such reasons do need to be provided and, moreover, that doing so does not represent an obviously straightforward normative task. And if I am right, then it cannot be said that it is easier to defend extending principles than limiting principles, all the way down.

IV. Comparative vs. Non-Comparative Free Speech Principles

The second distinction I want to analyse is between free speech principles that are comparative in nature and those that are non-comparative. Comparative principles (as the label suggests) are those that make assertions and/or prescriptions about free speech of a comparative nature. One subtype of comparative principles puts forward exemptions for speech from norms concerning protections and/or restrictions that apply to other freedoms. Consider the following three examples:

> A principle of freedom of speech asserts some range of protection for speech that goes beyond limitations on government interference with other activities.[66]

> [E]ven if speech causes harms which, if brought about by nonspeech conduct, could justify restricting that conduct, such harms do not justify restricting that speech.[67]

> Liberal free speech principles … demand that we accord a special status to 'speech', such that when it produces harm, it should be less liable to legal restriction than (similarly harmful) non-'speech'.[68]

A second subtype of comparative principles is exemplified by higher-level principles setting out the general quality or strength of justifications needed to warrant limitations on speech, as compared to the baseline of no such limitations or limitations on things that are not speech. Consider an example found in the work of Frederick Schauer—a free speech theorist who has also argued that the value(s) underlying the principle of free speech are not peculiar to speech but are in fact also promoted by a variety of other activities that are non-speech:[69]

66. Greenawalt, *supra* note 26 at 120.
67. Brison, *supra* note 4 at 317.
68. Simpson, *supra* note 11 at 457.
69. See Schauer, *supra* note 60.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

> Under a Free Speech Principle, any governmental action to achieve a goal . . . must provide a stronger justification when the attainment of that goal requires the restriction of speech than when no limitations on speech are employed.[70]

A third subtype of comparative principles is principles that make assertions about categories of speech the restriction of which requires stronger or weaker justifications compared to the restriction of other categories of speech. Consider the following political speech principle: Any governmental restriction of speech requires a stronger justification when the restriction of political speech is involved than when other categories of speech are involved.[71] Or consider the following hate speech principle: Any reason (based on values of free speech) that a public authority might have to refrain from restricting speech is weaker when the restriction targets hate speech in comparison to when the restriction targets other categories of speech.

A fourth subtype of comparative principles is principles that say something about which kinds of justifications for restricting speech are weaker or stronger than others. Consider the following harm/offence principle: Any justification for restricting speech that appeals to facts about the speech in question being offensive, disgusting, or shocking to others is weaker than one that appeals to facts about the direct and tangible harms of the speech because—when simply comparing the moral weight that should be given to the interest of a person in expressing their own opinion and the direct and tangible harm of their expression—the latter is weightier, but when simply comparing the moral weight that should be given to the interest a person has in expressing their own opinion and the feelings of another who is offended at their expression, the former is weightier.[72]

A fifth subtype of comparative principles is principles that assert justifications for restricting speech which in themselves involve comparing benefits and losses of alternative options vis-à-vis protection and restriction of speech. For example:

> When two important social values (such as liberty and equality) conflict, the optimal tradeoff or balance between them is that point at which further gains in one of the values would be outweighed by greater losses in the other. Freedom of expression would be better protected were there no legal constraints whatever on hate propaganda, while the equal status of minority groups would (arguably) be better safeguarded by legislation more restrictive than the hate propaganda law, hedged round as it is by its various safeguards. Somewhere between these extremes lies a balance point at which the greater protection for these groups afforded by more restrictive legislation would be outweighed by the greater impairment of expression, while the greater protection for expression afforded by more permissive legislation would be outweighed by the greater risk of discrimination.[73]

---

70. Schauer, *supra* note 2 at 7-8 [footnote omitted].
71. I take inspiration here from Sunstein, *supra* note 11.
72. This principle is loosely inspired by Mill's well-known comment on the religious bigot who offends other people. See Mill, *supra* note 5 at 152.
73. LW Sumner, *The Hateful and the Obscene: Studies in the Limits of Free Expression* (University of Toronto Press, 2004) at 62-63.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

Turning to non-comparative free speech principles, these are principles that make claims about free speech without drawing comparisons. Once again, there are several subtypes of non-comparative principles, some of which broadly correspond to the aforementioned subtypes of comparative principles. I shall offer three illustrations. First, consider the following non-comparative liberal free speech principle: We should accord a special status to 'speech' such that, even when it produces harm, it should not be liable to legal restriction unless the relevant harm reaches a certain threshold of harmfulness, at which point it should be so liable. Or:

> Most liberal defenders of free speech argue for a formal principle that concentrates on the protection of speech that is neutral about the ideas being expressed (up to the point where they instigate harm).[74]

Second, consider the following non-comparative justificatory free speech principle: Any governmental action to achieve a goal must provide a sufficiently strong justification when the attainment of that goal requires the restriction of speech.

Third, consider the following non-comparative hate speech principle: Any reason (based on values of free speech) that a public authority might have to refrain from restricting speech is surmountable when the restriction targets hate speech.

These are examples of non-comparative principles because they do not in themselves make or rely upon comparisons. The first example speaks to a threshold of harm, not merely more or less harm. The second example refers to a sufficiently strong justification needed to restrict speech and is silent on the justification required to restrict non-speech. The third example talks of reasons to refrain from restricting speech being surmountable when hate speech is in play and does not talk of whether the same holds for other categories of speech. What is more, there is no reason to assume that somehow each of these principles are implicitly comparative, as if they are only meaningful, comprehensible, or plausible if there are implied comparisons going on in the background. Arguably, each of these principles can be understood and potentially endorsed without reading into them hidden comparative assumptions.

The above should suffice to highlight the formal differences between comparative and non-comparative free speech principles. My main analysis of these differences boils down a series of arguments I wish to make in opposition to the belief or assumption some people may have that it is harder to prove, demonstrate, and motivate non-comparative than comparative principles; and conversely, that somehow defending comparative principles is the easier option.

First things first, why might someone think it is harder to prove, demonstrate, and motivate non-comparative free speech principles? Consider once again the non-comparative liberal free speech principle. Arguably, to adequately defend this principle would first require its champion to establish what the relevant

74. Warburton, *supra* note 2 at 76.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

'threshold of harmfulness' is. After all, if the threshold were too low, then the principle might not seem true or credible, but if the threshold were too high, then the principle could appear trivial in the sense of saying something entirely uncontroversial or uninteresting. But what is the correct threshold? And how will people come to reasonable consensus about it?

Similar difficulties would appear to dog the non-comparative justificatory free speech principle. Convincing people about the veracity and importance of this principle seems to necessitate furnishing a compelling explanation of what counts as a 'sufficiently strong justification'. But once again it seems a daunting task to come up with a standard of sufficiency that is neither too weak nor too strong. What is more, people are likely to have reasonable disagreements about what the phrase 'a sufficiently strong justification' ought to mean.

Likewise, non-comparative free speech principles which are bespoke to particular categories of speech are not easy to defend. Consider the non-comparative hate speech principle. What constitutes a 'surmountable' reason to refrain from restricting hate speech? And what grounds are there to accept the principle that *any* reason to refrain from restricting speech is surmountable when the speech in question is hate speech? Surely there could be things that reasonable people would consider insurmountable reasons to refrain from restricting speech even if the speech in question is hate speech. These are tough questions that would-be defenders of the non-comparative hate speech principle would face. And they are tough despite the fact that the principle does not make or rely upon comparisons with other categories of speech.

However, I believe it is wrongheaded to think such challenges are somehow absent or significantly less acute when considering relevant comparative free speech principles. Indeed, I would argue the level of difficulty is broadly similar. Consider once again the (comparative) liberal free speech principle which says we should accord a special status to 'speech', such that when it produces harm, it should be less liable to legal restriction than (similarly harmful) non-'speech'. The phrase 'similarly harmful' presupposes the possibility of making meaningful and accurate comparisons between the harms of speech and non-speech. But this is no mean feat. Harmfulness is not something that is easily identified and measured, unlike the temperature of water, for instance. In the case of an abstract property like harmfulness, there is need for the specification of a set of features of harmfulness that agents can use as criteria for assessing whether things are harmful and to what extent. What is more, for the comparison to be fair, roughly the same set of features (harmfulness criteria) must be applicable to all the things being assessed as potentially harmful. Yet specifying and coming to an agreement about a set of widely applicable features of harmfulness is not an easy task. It is certainly not easy when the harmfulness criteria must be applicable to two qualitatively different things that are putatively associated with constituting or causing harms, namely speech and non-speech; and potentially applicable to qualitatively different putative harms associated with speech and non-speech.

Similar difficulties confront the (comparative) justificatory free speech principle. Recall that it says any governmental action to achieve a goal must provide a

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

stronger justification when the attainment of that goal requires the restriction of speech than when no limitations on speech are employed. In order to be convinced of the truth of this principle and, what is more, in order for relevant agents to apply it, one first needs to understand what makes something 'a stronger justification'. Definitions of weaker and stronger justification need to be supplied and they must be applicable to both 'the restriction of speech' and 'no limitations on speech'. But once again, people are likely to have reasonable disagreements about what a 'stronger justification' ought to mean. As a result, proving, demonstrating, and motivating the comparative principle may not be an appreciably easier task than defending its non-comparative counterpart.

There is a further problematic ambiguity in the (comparative) justificatory free speech principle that deserves attention. This principle calls for a comparison between 'the restriction of speech' and 'no limitations on speech', but the second object of comparison (i.e., no limitations on speech)—*qua* baseline for comparison—is ambiguous between broader and narrower readings. In particular, the phrase 'no limitations on speech' could mean literally (i) the absence of limitations on speech, or instead it could mean (ii) the absence of limitations on speech yet also the presence of limitations on non-speech. In other words, strictly speaking, the phrase 'no limitations on speech' is compatible with describing a state of affairs (ii), in which there remain limitations on non-speech. However, if 'no limitations on speech' is understood to mean (ii), then the comparative principle is making a much more sweeping generalisation than at first appears. The problem is that the class of things plausibly defined as 'non-speech' is not merely extremely large but also includes many things that are extremely weighty and important and that arguably deserve at least as much if not more protection than freedom of speech. I have in mind things such as sexual freedom, reproductive freedom, freedom of movement, freedom of conscience, freedom of democratic participation, and freedom from torture, for example. If the principle is saying that any governmental action to achieve a goal must provide a stronger justification when the attainment of that goal requires the restriction of speech than when it requires the restriction of any one of these other basic liberties, then arguably many people will find the principle scarcely credible. They will say free speech is merely one among a number of equally important basic liberties characteristic of a liberal democratic society.[75] Of course, there are those who argue that free speech is in fact first among equals or *the* fundamental liberty, such as because they think it indispensable to all other liberties.[76] But those who make these arguments have to make them—there is nothing self-evident in what they argue.

Then again, if 'no limitations on speech' is interpreted to mean (i), then the (comparative) justificatory free speech principle could appear to be rather trivial, toothless, or redundant. For one thing, it probably becomes trivial to classify the principle as 'comparative', since the comparison is not between speech and other

75. For a list of basic liberties, see e.g. John Rawls, *A Theory of Justice* (Belknap Press, 1971) at 61.

76. See Eric Heinze, *The Most Human Right: Why Free Speech Is Everything* (MIT Press, 2022).

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

liberties but simply between the justifications required for limitations versus no limitations on speech. More importantly, it seems to be true of almost any desirable activity or conduct *qua* freedom that the government requires a stronger justification to limit that freedom than to refrain from limiting that freedom. Therefore, insofar as one could say there is a credible (comparative) justificatory free speech principle, one could just as easily say there is a credible (comparative) justificatory alcohol freedom principle, for instance. The latter principle might say that any governmental action to achieve a goal must provide a stronger justification when the attainment of that goal requires the restriction of alcohol consumption among adults than when no limitations on alcohol consumption among adults are employed. However, this would mean that the free speech principle no longer looks like a special principle.

Of course, some free speech theorists, such as Schauer, might be content to accept precisely that conclusion. But this does not mean others would be equally content. Part of the attraction and promise of (the idea of) the free speech principle would seem to be that it is a special principle that is not merely true but also distinctive or even unique, like free speech itself perhaps (recall the fifth desiderata introduction in Section II). Indeed, there is even a suspicion here that Schauer deliberately intends 'no limitations on speech' to carry the literal meaning (i) so as to render the free speech principle less special, something that chimes with his deeper scepticism towards the idea that free speech has underlying value(s) peculiar to it.

Perhaps one way around this would be to interpret the phrase 'no limitations on speech' to mean (iii) the absence of limitations on speech yet also the presence of limitations on *some* non-speech. On this reading, the principle is saying that any governmental action to achieve a goal must provide a stronger justification when the attainment of that goal requires the restriction of speech than when no limitations on speech are employed and when limitations on some forms of non-speech are employed. If interpretation or meaning (ii) risks turning the (comparative) justificatory free speech principle into an over-generalisation that fails to disaggregate forms of non-speech, then meaning (iii) seems to capture the intuition that not all non-speech and not all limitations on non-speech are qualitatively the same. At the same time, like meaning (ii), meaning (iii) lends the principle greater epistemic ambition and analytical bite—and greater distinctiveness—than under meaning (i). However, the realisation of that ambition and bite—and distinctiveness—is dependent on the defender of the principle, when interpreted as meaning (iii), being able to furnish a persuasive account of which freedoms are included within the scope of the principle and which are not. If meaning (iii) included the presence of limitations on the freedom of alcohol consumption or pet ownership, say, then the principle might ring true, but what about other freedoms? Which are the freedoms that are ranked alongside speech in terms of requiring the same strength of justifications for restrictions (i.e., basic liberties) and which are the freedoms that rank below speech and so require less strong justification for restrictions? Surely there is nothing easy—and nothing easier—about coming up with a list of freedoms that everyone could reasonably accept as

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

being covered under the scope of the principle interpreted in this more nuanced way.

A similar story can be told about (comparative) free speech principles which are bespoke to particular categories of speech. Consider once again the (comparative) hate speech principle: Any reason (based on values of free speech) that a public authority might have to refrain from restricting speech is weaker when the restriction targets hate speech in comparison to when the restriction targets other categories of speech. Surely it would be extremely hard to defend this principle the moment one begins to compare hate speech with other extremely harmful categories of speech. After all, what about the category of child pornography? Nobody could seriously imagine that any reason a public authority might have to refrain from restricting speech is weaker when the restriction targets hate speech in comparison to when the restriction targets child pornography. Consequently, to achieve credibility, the (comparative) hate speech principle will almost certainly need to be watered down with the qualification 'in comparison to when the restriction targets *some* other categories of speech'. But which other categories? There is nothing easy—or easier—about answering that question.

V. Monistic vs. Pluralistic Principles

The final distinction I want to analyse is between monistic and pluralistic free speech principles. Starting with the former, such principles set out some morally relevant feature that they assert is the exclusive or only feature of free speech in some particular sense specified by the principles. One subtype of monistic principles is exemplified by principles that advance one value as being either (i) the only justification for protecting free speech, (ii) the fundamental justification upon which other, derivative justifications depend, or (iii) the justification of most importance in some sense different to (ii), such as the justification that is of primary concern or operative for the most important free speech protections (e.g., categories of speech). Consider three examples of such principles:

> [T]he values supported or functions performed by protected speech result from that speech being a manifestation of individual freedom and choice.[77]

> [I]f one had to identify the single value that was uppermost in the minds of the persons who drafted and ratified the First Amendment, [the] checking value would be the most likely candidate.[78]

> [A] thinker-based foundation undergirds the most important free speech protections.[79]

77. C Edwin Baker, "Commercial Speech: A Problem in the Theory of Freedom" (1976) 62 Iowa L Rev 1 at 3.
78. Blasi, *supra* note 22 at 527.
79. Shiffrin, *supra* note 17 at 284.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

Another subtype of monistic principles is exemplified by a principle that sets forth the only fitting or rightful purpose for restricting speech by the lights of the principle (one necessary condition). Instances of this subtype of monistic principles often take their cue, whether in substance or style, from Mill's harm principle:

> That the only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others.[80]

Taking inspiration from both Mill's harm principle and his famous corn-dealers example,[81] some scholars interpret Mill as envisaging an absolute right to liberty of thought but only a limited right to freedom of expression, where this right is limited by the rightful purpose of harm prevention.[82] This interpretation seems to support the following Millian free speech principle: That if the only purpose for which power can be rightfully exercised over any member of a civilised community, against their will, is to prevent harm to others, and if certain limitations on speech can count as the exercise of such power, then it follows that the only purpose for which those limitations on speech can be rightfully exercised is to prevent harm to others. (Of course, not every scholar takes the exegetical view that, for Mill, people have a limited right to freedom of expression. Others argue instead that, for Mill, people do have an absolute right to freedom of expression, but a limited right to freedom of action, including no right to commit acts of using words in particular contexts with a view to inciting angry mobs to violence.)[83]

An alternative way of reading what Mill does, or should, say about the limits of free speech would be to adopt a nuanced interpretation of the argument he develops in defence of free speech in chapter 2 of *On Liberty*. Consider the nuanced Millian free speech principle: That the only circumstances in which official suppression of opinions can be rightfully carried out is when it can be demonstrated with good evidence that suppressing the relevant opinions prevents harm to others and does *not* have baneful consequences for the permanent interests of humans as progressive beings, even if the opinion being suppressed is true, even if it is false, and even if it is part true/part false. Of course, there will be those who think that no such circumstances exist in practice, but nevertheless some people may take a different view. What is more, affirming the nuanced Millian free speech principle does not commit one to thinking it is a pluralistic principle if the harm prevention component is also understood in terms of the master value of the progressive being.

In addition to this, some contemporary writers argue that Mill erred in focusing almost solely on the epistemic gains that attend to refraining from restricting

80. Mill, *supra* note 5 at 78.
81. *Ibid* at 123.
82. See Jonathan Riley, "J.S. Mill's Doctrine of Freedom of Expression" (2005) 17:2 Utilitas 147 at 159; John Skorupski, *John Stuart Mill* (Routledge, 1989) at 374.
83. See KC O'Rourke, *John Stuart Mill and Freedom of Expression: The genesis of a theory* (Routledge, 2001) at 153; Daniel Jacobson, "Mill on Liberty, Speech and the Free Society" (2000) 29:3 Philosophy & Public Affairs 276 at 280.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

speech, such as gains in truth discovery and knowledge acquisition, and paid insufficient heed to the non-epistemic losses that can accrue from not restricting speech in some circumstances, such as losses in security, dignity, and freedom from oppression and subordination.[84] To counter this imbalance, Schauer proffers an alternative principle:

> The post-Millian calculus ... suggests that an institution or practice of suppression, for some category of utterances within some domain of speakers and listeners (or writers and readers), can be justified only when (but not always when) it is predicted that the consequential losses from the spread of false opinions that might be accepted and acted on despite their falsity will be greater than the consequential gains that will come from the discovery of previously unknown truths and the increase in knowledge that is the corollary of that discovery.[85]

At first glance, because the post-Millian calculus involves the weighing up of both epistemic and non-epistemic gains and losses, there is one sense in which it is a pluralistic free speech principle, based on different sorts of values (the first subtype). Yet in another sense (the second subtype) it remains a monistic free speech principle, based on the relevant fact of positing one necessary condition. This is because it is predicated on the idea that restricting speech can be justified only based on one particular sort of balancing act or calculus.[86]

Turning to pluralistic free speech principles, this sort of principle either advances one thing that it deems to be a morally relevant feature of free speech without asserting or implying that it is the only important feature, or else advances several things at the same time. One subtype of pluralistic free speech principles includes principles that set forth two or more features capable of supporting (including functional arguments) free speech protections without further claiming these are the only such features. Consider three examples of such principles:

> The right to freedom of expression ... derives from the widely accepted premise of Western thought that the proper end of man is the realization of his character and potentialities as a human being.... In the traditional theory, freedom of expression is not only an individual but a social good. It is, to begin with, the best process for advancing knowledge and discovering truth.... The third main function of a system of freedom of expression is to provide for participation in decision-making through a process of open discussion which is available to all members of the community.[87]

> The values sought by society in protecting the right to freedom of expression may be grouped into four broad categories. Maintenance of a system of free expression is necessary (1) as assuring individual self-fulfilment, (2) as a means of attaining the truth, (3) as a method of securing participation by the members of the society in

84. See Brown, *supra* note 46 at 225-26.
85. Frederick Schauer, "Social Epistemology, Holocaust Denial, and the Post-Millian Calculus" in Herz & Molnar, *supra* note 24, 129 at 138
86. For an application of this sort of balancing act or calculus to various examples of hate speech, see Brown, *supra* note 46 at ch 8.
87. Emerson, *supra* note 16 at 879, 881, 882.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

> social, including political, decision-making, and (4) as maintaining the balance between stability and change in the society.[88]

> Freedom of speech should be valued for many reasons—not only liberty, self-realization, freedom, and autonomy but also truth, combating injustice, adaptation to change, democracy, equality, association, freedom of thought, and even order.[89]

Another subtype of pluralistic principles is exemplified by limiting principles that establish two or more things capable of justifying limitations on speech. Consider this example:

> What does the state, acting on behalf of society as a whole, owe to citizens when it comes to regulating speech or other modes of expression? Some people believe that in answering this question it makes a positive difference whether or not the speech in question is insulting, degrading, defaming, negatively stereotyping or inciting hatred, discrimination or violence against people in virtue of their race, ethnicity, nationality, religion, sexual orientation, disability, gender identity, for example; and that it makes a positive difference because such speech implicates issues of harm, dignity, security, healthy cultural dialogue, democracy, and legitimacy, to name just a handful of relevant issues.[90]

My analysis of the differences between monistic and pluralistic free speech principles is comprised of three interconnected points. My first and central point is that, other things remaining equal, it *is* harder to prove, demonstrate, and motivate monistic than pluralistic free speech principles; and conversely, that defending pluralistic principles is the easier option. For example, it is harder to defend a principle that asserts something is the only, the fundamental, or the primary morally relevant feature supporting free speech protections than to defend a principle that asserts simply that some feature supports free speech protections and is perhaps one among many such features. Likewise, it is harder to defend a principle that asserts something is the only rightful justification for restricting speech than to defend a principle that asserts simply that something is one rightful justification and perhaps one among others. The reason for this is that the phrase 'the only' is an exhaustive quantifier meaning that it implicitly answers the question 'How many?' with the answer 'just one'. It also sets up a necessary condition, namely that something x must be present in order for free speech protections to be supported, or something x must be present in order for limitations on speech to be rightfully justified. By contrast, the phrase 'one thing' is an existential quantifier meaning that it answers the question 'How many?' with the answer 'there is at least one (and perhaps others)'. This does not set up a necessary condition (provided that the principle does not assert a finite list of features such that their disjunction, as in, this or this or this, is a necessary condition). Instead, it might

---

88. *Ibid* at 878-79.
89. Steven Shiffrin, "Dissent, Democratic Participation, and First Amendment Methodology" (2011) 97:3 Va L Rev 559 at 559.
90. Alexander Brown, "What is Hate Speech? Part 1: The Myth of Hate" (2017) 36:4 L & Philosophy 419 at 419-20.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

create a weak sufficient condition, namely that if something x is present, then free speech protections can be supported, or that if something x is present, then restricting speech can be rightfully justified. Furthermore, the exhaustive quantifier 'the only' entails there are no other features, which is a negative claim. And it can be harder to prove a negative than a positive. In other words, it is incumbent on monistic principles to provide an account or explanation of both (i) why a particular feature matters to free speech in some relevant way, and (ii) why nothing else matters or nothing else matters as much in the relevant way. By contrast, pluralistic principles need only explain (i).

My second point is designed to cast doubt on a further putative explanation that might be offered as to why it is harder to prove, demonstrate, and motivate monistic free speech principles; and conversely that defending pluralistic principles is the easier option. The explanation appeals to the distinction between all-things-considered and prima facie support/justification. It is essentially that monistic principles typically invoke all-things-considered support/justification and pluralistic principles usually invoke prima facie support/justification.

However, I believe it is wrong to assume the distinction between all-things-considered and prima facie support/justification holds the key to explaining why it would be easier to defend pluralistic free speech principles. To fully understand this point, it is necessary to dwell a little further on the former distinction. Consider the following, more detailed explanation of the distinction.

> One kind of warrant is concerned with whether or not a law/regulation/code is *prima facie* justified, authorized, sanctioned, supported, or rendered permissible by the contribution it makes to a given normatively relevant feature, such as a right, interest, [duty, public good,] or value. When a principle specifies whether a law/regulation/code is *prima facie* warranted or unwarranted with reference to a given normatively relevant feature—meaning that the principle's verdict holds unless it is overridden or trumped by another principle which itself may highlight a different normatively relevant feature—I shall call this *narrow warrant* or *N-warrant*, for short. Of course, the larger the number of relevant principles, the lower the chances that any law/regulation/code will be N-warranted by each and every relevant principle. In the main, legislatures, courts, and regulators will be called upon to decide between a law/regulation/code that is N-warranted by one or more principles but also N-unwarranted by one or more principles. A second kind of warrant is tailored to addressing precisely these sorts of dilemmas. It requires overarching determinations of whether a law/regulation/code is warranted or unwarranted based on every relevant principle. I shall call this *all principles considered warrant* or *overall warrant*, that is, *O-warrant*, for short.[91]

The academic literature on free speech contains numerous limiting free speech principles that set forth prima facie justifications or N-warrants of limitations on speech, but which fall short of providing all-things-considered justifications or O-warrants. Consider two examples:

91. Brown, *supra* note 46 at 3 [emphasis in original].

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

> [T]he Principle of Health [says] that legalistic constraints on uses of hate speech are (N-)warranted if they function to protect people from severe damage to their psychological or physiological health. . . . [I]n order to pass muster under the Principle of Health the relevant damage to health must be reckonable (amenable to measurement using established techniques and methods in the social sciences), assignable (attributable to particular individuals and not merely to an entire group of people en masse), and severe (beyond the level of mild irritation or feelings of offense). . . . [T]he Principle of Health is concerned with N-warrant. Would a given piece of hate speech law be *prima facie* warranted or unwarranted by *this* particular normatively relevant feature?. . . [T]he question of overall warrant is a different matter. In order to evaluate overall warrant, legislators and judges would also need to consider a range of other principles that invoke practical as well as moral considerations, such as whether a particular law would be effective in preventing the relevant harms, would be the least restrictive method of achieving that result, and would not have serious unintended consequences for free speech.[92]

> [T]he moral right to freedom of expression, properly interpreted, does not protect speech that incites clear violations of others' moral rights. Instead . . . there is an enforceable moral duty to refrain from such incitement, a duty that shapes and limits the moral right to free speech itself. . . . [However], 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 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. . . . [O]ther familiar arguments in the free-speech literature—for example, 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.[93]

Reflecting on this distinction between all-things-considered and prima facie justifications of limitations on speech, or O-warrant and N-warrant, some people might assume that it would be easier to defend a limiting free speech principle that sets out a prima facie justification or N-warrant of certain limitations on speech than to defend a principle which asserts an all-things-considered justification or O-warrant of the relevant limitations. That is because the scale and demandingness of all-things-considered justification or O-warrant is much greater in the sense that it needs to cover more and do more. Specifically, a limiting free speech principle which asserts an all-things-considered justification or O-warrant of certain limitations on speech must not only cover all relevant considerations but also make the case for limiting speech even in the face of counter-considerations.

However, there is a flaw in this line of explanation. For, on closer examination, it is not *substantially* easier to defend a limiting free speech principle that sets out a prima facie justification or N-warrant of certain limitations on speech

92. *Ibid* at 51.
93. Howard, *supra* note 2 at 210-11 [emphasis removed].

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

than to defend a principle which asserts an all-things-considered justification or O-warrant of the relevant limitations. How so? Consider the two principles quoted above which appeal to, respectively, protecting people from severe damage to their psychological or physiological health and the duty to refrain from inciting clear violations of others' moral rights. Even if one accepts that these principles merely set out prima facie justifications or N-warrants of limitations on speech, this does not somehow make them easy to defend or even substantially easier to defend. After all, the defender of these principles has the very significant challenge of showing why the relevant moral features (protection of health and the duty to refrain from incitement) would indeed justify limitations on speech even in a prima facie way. Given that threatening people with prison sentences, say, for breaking hate speech laws might lead people to self-censor and repress their true expressive feelings or else to end up behind bars for years in a way that could be damaging to their psychological or physiological health, why does protection of health justify hate speech bans even in a prima facie way? That is a difficult question. And, given the many descriptive and moral differences between inciting clear violations of others' moral rights and actually perpetrating clear violations of others' moral rights, why should we believe there is a duty to refrain from incitement and, moreover, why should we suppose this duty justifies incitement laws even in a prima facie way? Now I do not mean to suggest these are unanswerable questions. Rather, I make the different, more nuanced point that these questions are very thorny normative questions, so it is hard to understand the sense in which these questions are substantially easier to answer than the questions faced by defenders of pluralistic free speech principles.

My third point relates to the definition of harm within free speech principles (or those principles that invoke the notion of harm). Some people might try to undermine my claim that it is harder to defend monistic principles than pluralistic principles by relying on a particular account of harm. Specifically, they might imagine it easier to defend harm-based monistic principles by embracing what I shall call a dependent account of harm. To properly assess this putative counter-argument, I first need to explain the distinction between independent and dependent accounts of harm.

An independent account of harm is one that explains what harm is and identifies relevant harms independently of the free speech principle, which is to say that an independent account treats the question 'What is harm?' as logically prior to the question 'When is it legitimate to restrict speech?'. In other words, the independent account of harm informs the theory of free speech and not the other way around. However, coming up with a credible independent account of harm is not easy, especially because there is no fixed point of reference for saying what the purpose of defining harm is supposed to be. The independent account faces countless challenging questions. For example, should the idea of harm include not only tangible harms such as bodily injuries but also psychological injuries such as severe emotional distress? What about indirect harm, such as an increased susceptibility to anxiety and depression? More generally, is risk of harm itself a form of harm? Should feelings of deep offence qualify as psychological harms?

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

Can things other than bodily and psychological injuries also count as harms? What about dignitary injury? What is dignitary injury? Does it include being the target of an illocutionary speech act of ranking one as inferior, for example? And what about harms to a person's interests? Does any setback to a person's interests count as harm in the relevant sense, or is Joel Feinberg correct to insist that "only setbacks of interests that are wrongs, and wrongs that are setbacks to interest, are to count as harms in the appropriate sense"?[94] And, what are wrongs? Does the class of wrongs include accumulative or aggregative harms? If not, does this show there can be harms that are not wrongs after all?

Since some of the limiting free speech principles in the literature take inspiration from Mill, one option is to seek answers in his theory of harm. However, once again his theory furnishes as many questions as answers. For example, Mill's position was that conduct is harmful if it injures another person's interests, or to be more exact, "certain interests, which ... ought to be considered as rights."[95] But which are these interests? On my reading, they are what Mill called "permanent interests."[96] After all, Mill argued not only that utility is "the ultimate appeal on all ethical questions" but also that this must be understood as "utility in the largest sense, grounded on the permanent interests of man as a progressive being."[97] Then again, what are these permanent interests? Roughly speaking, for Mill, the relevant interests must be universal in nature and related to the dignity of humans as thinking beings and as beings capable of individuality and idiosyncrasy.[98] But which are these interests?

I do not mean to suggest such questions can never be adequately answered, merely that it is not easy to answer them—or not easier. Moreover, it is one thing answering these general questions about harm in a way that is coherent and likely to enjoy reasonable agreement, it is quite another to answer specific questions about how harm defined in a certain way as it relates to particular categories of speech and even to particular instances of those categories. Consider hate speech. What harms are associated with hate speech? Do all forms of hate speech involve the same kinds of harm or are different forms of hate speech associated with different kinds of harm? For example, do racial microaggressions involve the same quality and quantity of harm as incitement to racial discrimination? Does online hate speech inflict the same quality and quantity of harm as offline hate speech? At what point exactly does the quality and quantity of harm cross a threshold such that there is a prima facie justification for restricting hate speech? Once again, there might be good answers to these questions, but finding them is not easy—or easier.

Rather than having to keep answering these questions—or in the event that one reaches a terminus where one is simply no longer able to provide persuasive

94. Joel Feinberg, *The Moral Limits of the Criminal Law, Volume One: Harm to Others* (Oxford University Press, 1984) at 36.
95. Mill, *supra* note 5 at 143.
96. *Ibid* at 79.
97. *Ibid* [footnote omitted].
98. See *ibid* at chs 2-3.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

answers—one could embrace a dependent account of harm. According to a dependent account, 'harm' is a label we give to something for which there is a pro tanto reason to restrict speech that constitutes or produces it. In other words, a dependent account draws a very tight conceptual connection between the justifications for restricting speech and what qualifies as harm under the relevant limiting free speech principles, treating the relevant harm as simply that which is identified by good justifications for limitations on speech.[99]

To give a concrete illustration, a dependent account of harm would deal with the dilemmas caused by hate speech as follows. It seems there is at least prima facie justification to legally restrict certain hate speech if that hate speech causes emotional damage, contributes to a climate of hatred, inflicts dignitary injury, involves misrecognition of identity, undermines public assurances of civic dignity, and reduces the political legitimacy of the state.[100] Then again, at first glance, it is not clear that going ahead with the restriction would satisfy the nuanced Millian free speech principle that says the only purpose for which certain limitations on speech can be rightfully exercised is to prevent harm to others. That is because it is unclear whether the aforementioned phenomena count as harms. One solution therefore would be to classify these phenomena as harms simply in virtue of the fact there seems to be a prima facie justification for limiting hate speech that produces them.[101]

It appears as though embracing the dependent account of harm would make it easier to defend a harm-based monistic free speech principle. After all, when presented with a putative counter-example of phenomena that seem to justify restricting speech (e.g., dignitary injury), the current solution is simply to swallow up the example by counting the relevant phenomena as harms purely by dint of the fact that their occurrence seems to justify restricting speech. The apparent advantage of the dependent account of harm is that one can retain both the notion that harm prevention provides a prima facie justification for restricting speech and the notion that harm prevention is the only rightful purpose for restricting speech. This is because the dependent account simply defines 'harm' to mean anything that is morally significant enough to constitute a pro tanto reason for restricting speech. Another advantage might be that in ordinary language the term 'harm' is a catchall that seems to cover a whole range of things going badly for a person or entity, meaning a hurt, injury, damage, an evil, wrong, loss, and so on,

99. Some people might respond to this by saying I have overlooked some kind of hybrid account of harm that takes into consideration pro tanto reasons for limiting speech but is not slavish to these reasons such that something could be defined as harmful partly because of these reasons, but at the same time these reasons are not the only features defining *harm*, meaning something could be harmful even in the absence of these reasons. Such an account might be possible but is not one I have space to explore in this paper. My main focus here is on the deeper distinction between monistic and pluralistic free speech principles. I use the distinction between dependent and independent accounts of harm as a way of teasing out the comparative difficulty of defending monistic and pluralistic principles.

100. See Brown, *supra* note 46.

101. Note, I am not suggesting Mill embraced the dependent account of harm.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

but the dependent account reconceptualises harm so as to guarantee it can be put to service in a harm-based free speech principle. By contrast, the independent account of harm must continue to grapple with the capacious, sometimes irrelevant, and often incoherent ordinary language uses of the term 'harm' with the aim of eventually moulding the idea into something that could successfully work inside a harm-based free speech principle but without the guarantee of success that it will.

However, I believe these advantages are not all they are cracked up to be and are outweighed by the disadvantages of the dependent account of harm. For one thing, the dependent account ends up trivialising or removing substance from the notion that harm prevention provides a prima facie justification for restricting speech. For, under the dependent account, this is no longer a notion that says something informative about harm; instead, this notion becomes a kind of tautology, true simply by virtue of the meaning given to the term 'harm' under the dependent account. Putting this another way, it seems intuitive to think that we might want to begin with the question 'Is it legitimate to restrict speech?' and then proceed to answer it in a substantive way as follows, 'Yes, when the speech is harmful'. The answer is substantive because it promises to genuinely advance our aim of answering the initial question, as in, to provide illumination. The answer might not in the end realise that promise—because of the difficulty of explaining what counts as harm—but at least the promise is there. However, if one defines harm as being simply whatever is identified in any good answer to the initial question, then this does not genuinely advance our aim of answering the initial question because it does not tell us what a good answer is.

Of course, one could seek genuine advancement by developing an independent account of what constitutes a prima facie justification for restricting speech. But then again, one might have expected the notion of harm prevention to be part of the answer. Moreover, if the dependent account of harm is ultimately reliant for its success on developing an independent account of what constitutes a prima facie justification for restricting speech, then arguably it would no longer be plausible to describe the dependent account of harm as the easier option. Indeed, it would appear to be an even harder option because of the reliance on developing an independent account of justification in the relevant sense. More importantly, this project takes us beyond a monistic limiting free speech principle towards a pluralistic principle in the event the justification relies on two or more fundamental values.

Furthermore, it seems intuitive to think there could be legitimate reasons to restrict speech that have nothing to do with the harmfulness of speech under a strict definition of harm as bodily or psychological injury, and that could apply to harmless speech, namely categories of regulatable yet harmless speech (under a strict definition of harm), such as hate speech that constitutes dignitary injury or misrecognition of identity. I believe this technical idea of regulatable yet harmless speech can have an important role to play in justifying particular kinds of hate speech restrictions, whether that is Holocaust denial laws, group defamation

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

laws, or even some delicts.[102] But if one embraces the dependent account of harm, then this is no longer a possibility. The regulatable becomes by definition the harmful and the harmful the regulatable.

The converse reason not to adopt the dependent account of harm might be the usefulness of retaining a concept of harmful speech that is reasonably narrow and picks out a relatively limited set of phenomena, such as damage to a person's physical body or mental well-being of the sort that can be grounded in evidence supplied by medical professionals. Such a concept might also play an important role in justifying particular kinds of hate speech restrictions, including some civil torts.[103]

In other words, if the concept of harm is inflated to occupy the entire justificatory landscape (as per the dependent account), then it could be harder to individuate and separate different sorts of justifications for restricting speech. An all-encompassing category of harm might undercut nuanced arguments about which particular justifications are better suited to some hate speech laws than others, based on the distinctive qualities of both the justifications and the hate speech laws.[104] For example, saying that harm is simply anything that is morally significant enough to constitute a pro tanto reason for restricting speech smacks of one scale of moral significance and one form of justification. However, if one says instead that different harms are very different in nature to each other and different again from non-harms and that the differences really matter when it comes to justifying given limitations on speech, then this facilitates concluding not only that different harms can justify different kinds of limitations on speech but also that different harms might be involved in different varieties of moral significance and even different forms of justification itself.

Moreover, defining the term 'harm' as 'anything that is morally significant enough to constitute a pro tanto reason for restricting speech' risks eliding differences between the types of empirical evidence needed to support or ground different sorts of harm-based justifications for restricting speech. After all, the sorts of empirical evidence needed to ground (medical) hypotheses that certain forms of hate speech can cause medium- to long-term psychological and physiological harms[105] is different to the sorts of evidence needed to ground (sociolinguistic) hypotheses that certain forms of hate speech can successfully perform acts of subordination (e.g., ranking others as inferior),[106] and once again different to the sorts of evidence needed to ground (sociological) hypotheses that certain forms of hate speech can contribute to a climate of hatred that predicts an increased risk of acts of discrimination or violence[107] and that the contribution is typically

102. See Brown, *supra* note 46 at ch 3.6, ch 6.2; Alexander Brown, "Retheorizing Actionable Injuries in Civil Lawsuits Involving Targeted Hate Speech: Hate Speech as Degradation and Humiliation" (2018) 9:1 Alabama Civil Rights & Civil Liberties L Rev 1.
103. See Brown, *supra* note 102; Matsuda, *supra* note 46.
104. See Brown, *supra* note 46 at 316.
105. *Ibid* at ch 3.1.
106. *Ibid* at ch 3.4.
107. *Ibid* at ch 3.3.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

increased when the speakers are politicians.[108] If differences between these sorts of evidential grounds lend different levels and kinds of epistemic credibility to given claims about the effects of certain speech, then it would seem an advantage (i.e., more accurate) to reflect this variety in definitions of harm. Arguably, this variety is more easily reflected in independent than dependent accounts of harm. For, in the case of dependent accounts, the notion of epistemic credibility is tied to the idea of what could potentially justify limitations on speech. And this might arbitrarily narrow what we understand epistemic credibility to be. There could be a tendency to think that epistemic credibility means 'what can be proven in a court of law'. That notion might have some useful applications, of course, but as a universal assumption about the concept of harm, it is arguably under-inclusive. Ordinary people in their everyday lives might sometimes reasonably employ an understanding of harmful speech that does not rely on epistemic credibility defined in the aforementioned narrow sense.

More generally, as it stands, in much of the academic literature on free speech and in much of the relevant case law to which it is sensitive, the aim or purpose of harm prevention is seen as setting a high bar for justifying limitations on speech. And scholars and legal professionals who are attracted to harm-based free speech principles are typically so attracted because harm sets a high bar. But this high bar might not be appropriate for ordinary people in their everyday lives who wish to talk about harmful speech. On the other hand, some scholars and legal professionals may foresee a risk that the concept of harmfulness ends up being accidentally weakened as a consequence of tethering it to the broad notion of pro tanto reasons for limitations on speech. This fear might be present among those who are attracted to monistic harm-based limiting free speech principles. They think harm sets a high bar for limitations on speech and that it ought to be the only bar (partly because they think it should never be easy to justify limitations on speech). However, the dependent account of harm could potentially lower the bar by expanding the meaning of the term 'harm' to encompass anything that could provide a prima facie justification for restricting speech. There may be innovations in legal scholarship and legal practice that accept features like 'harm to democracy', say, as prima facie justifications for restricting speech, and these innovations might turn the concept of harm into a lower threshold for justifying limitations on speech owing to the dependent account.

Finally, the dependent account of harm only appears to provide a solution to the incommensurability problem: that to talk of adding together different gains from restricting speech and weighing the total gains against the total losses of doing so is meaningless because the phenomena on the scales are qualitatively different and lack a shared basis for comparison with each other. Describing all of the different gains as 'harm preventions' and the relevant losses as 'harms to free speech' might seem to address this problem by making the different desiderata all of a piece—as so many 'harms'—but in truth the dependent account of

108. See also Brown & Sinclair, *supra* note 65 at ch 7.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

harm only superficially solves the incommensurability problem. It is just a linguistic move to label as 'harm' all things that are relevant to moral assessments of proposed limitations on speech. It remains likely that the various 'harm preventions' (e.g., preventing speech that inflicts severe emotional distress, preventing speech that contributes to a climate of hatred, preventing speech that constitutes dignitary injury or misrecognition of identity) will be qualitatively distinct and incommensurable, both with each other and with the relevant losses that attend limitations on speech (e.g., undermining truth discovery, violations of formal autonomy, diminishing democratic self-government).

The moral of this story is that the dependent account of harm is not a magic bullet that automatically makes it easier to defend harm-based monistic free speech principles. On the contrary, it makes it either no easier or harder to defend such principles (a liability)—if by 'easier' one means easier to defend as true, as applicable, and as capable of commanding a reasonable consensus. I believe similar points may also apply to free speech principles that invoke morally relevant features other than harm. Consider an autonomy-based monistic free speech principle that asserts that safeguarding autonomy is the only proper support for free speech protections and also the only justification there can be for rightful limitations on speech. But what is autonomy? Is it a matter of the freedom to think and choose without undue influence from external sources? What counts as undue influence? If the champion of autonomy-based free speech principles embraced a dependent account of autonomy, then they would say simply that undue influence is the sort of influence that is rightfully restricted. However, this dependent account of autonomy seems every bit as unsatisfactory as the dependent account of harm, and for the same reasons.

VI. Implications

I have argued that the philosophical free speech principle is not in fact a monolith but rather a cluster of principles. I have also critically assessed three main differences or distinctions between such principles: (i) extending vs. limiting principles; (ii) comparative vs. non-comparative principles; and (iii) monistic vs. pluralistic principles. No doubt each of these distinctions stands in need of further investigation and will face demands to be further defended. For each there will be sceptics who say that the distinction collapses. Some people might think that under closer scrutiny all free speech principles are in fact comparative, for example. Furthermore, it is not always straightforward to categorise principles actually defended by free speech scholars. Steven Shiffrin, for example, defends a free speech principle framed around the value of dissent but simultaneously contends that a focus on dissent does not "liquidate" concern for a long list of other familiar free speech values.[109] The dissent free speech principle has the appearance of a

109. Steven H Shiffrin, *The First Amendment, Democracy, and Romance* (Harvard University Press, 1990) at 167.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

monistic principle but is clarified as being a pluralistic principle. However, this article is primarily aimed at those people who intuitively accept, and can for the most part apply, these three distinctions and who, importantly, are minded to believe that certain free speech principles are easier or more difficult to defend depending on which sides of the distinctions the principles fall.

Clearly there are a large number of permutations (types of free speech principles) produced by these three distinctions: limiting non-comparative pluralistic principles, limiting comparative pluralistic principles, limiting non-comparative monistic principles, and so on. I do not have space here to fully investigate and evaluate all permutations. An ambition of this article is to inspire a research agenda, namely, to motivate the need for future research looking into the inter-relationships between these distinctions, as well as into which permutations appear most promising, especially when it comes to illuminating general debates around justifying free speech and its limits, and more specific debates around particular examples of controversial categories of speech like hate speech.

However, there are three immediate general implications I do wish to tease out in this article. The first has to do with conceptual complexity. Although my aim has been to clear up certain conflations and confusions in how the phrase 'the free speech principle' is used in the academic literature and to provide a conceptual map of the terrain of free speech principles, this does not mean the map itself is a simple one. On the contrary, the aforementioned distinctions may sometimes track or overlap each other but not always: They sometimes demarcate the same principles into the same groups, but they sometimes put them into different groups. With three cross-cutting distinctions, there are many permutations, such that different principles can have different combinations of qualities. For example, what Schauer calls "the post-Millian calculus" (Section V) is in one sense a limiting principle, a comparative principle, and a monistic principle.[110] It is a limiting principle because it talks about a particular justification for restricting speech; it is a comparative principle because the justification it sets forth involves comparing (for the purposes of a balancing exercise) consequential gains and losses of restricting speech; and it is a monistic principle because it asserts that the relevant justification (a balancing exercise) is the only legitimate justification for restricting speech. However, an alternative free speech principle could just as easily be an extending principle, a non-comparative principle, and a pluralistic principle. Other free speech principles could exemplify yet further combinations of qualities. Thus, the conceptual map I am drawing is potentially extremely complex.

The second implication I want to tease out concerns the logical relationship between free speech principles and free speech policy dilemmas. The key point is that free speech is a matter of principle as well as policy, which is to say there are such things as philosophical free speech principles which invoke morally relevant features such as interests, rights, duties, public goods, and fundamental values,

110. Schauer, *supra* note 85 at 138.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

but there are also doctrinal free speech principles (policies) grounded in bodies of law, and also specific free speech policy dilemmas concerning whether it is the right thing to do to enact, enforce, and apply a certain law or regulation that restricts some category of speech within some domain or context and for some speakers and listeners.[111] What is the logical relationship between these principles, policies, policy dilemmas, and adjudications?

One way to think about the relationship is that the direction of analysis runs from principles to policies to adjudications. For example, a philosophical free speech principle might inform or justify some doctrinal free speech principle (e.g., the First Amendment) or principles (e.g., various specific First Amendment doctrines such as the strict scrutiny test or the doctrine of content and viewpoint neutrality), and together both philosophical and doctrinal free speech principles might guide specific policy decisions, such as whether or not to uphold or strike down a particular hate speech law. However, once it has been recognised that the philosophical free speech principle is not in fact a monolith but rather a cluster of principles, it might be asked why the direction of analysis should run from principles to policies. Why should it not run in the reverse direction from policies to principles? For example, a government might begin with some dilemma over whether to pursue policy x or y (e.g., upholding a hate speech law versus striking it down; or enacting a hate speech law versus promoting public counter-speech, engaging in official counter-speech, launching education campaigns, funding victim support programmes, and other extralegal measures). It could then figure out which free speech principle suggests x and which suggests y. Having decided for some independent reason (e.g., sending a message to certain voters that it is listening to them) to choose policy x over y, the government could then seek to justify its decision by pointing out that x is supported by a certain free speech principle.

This might be an accurate assessment of how some policymaking is done—namely, the government or court internally decides what it wants to do and then, when speaking to parliament, the media, and society at large, it retrospectively defends the decision by invoking some sort of principle—but it is not an appealing picture (or so says a philosopher).[112] If there is a bedrock principle of legal and political philosophy and normative jurisprudence in particular, it is that philosophical principles should be used to critically assess, and ultimately justify or tell against, doctrinal principles, policies, and adjudications in a meaningful way. That is not to say the policy landscape should not shape and inform the principles. After all, if principles are to function as policy guides, they need to reflect to some extent what the policy dilemmas and policy options actually are.[113] However, the

111. See also Ronald Dworkin, *A Matter of Principle* (Harvard University Press, 1985).

112. It might not even be an accurate description of how policymaking has been done in the field of hate speech law in some countries. For an alternative vision, see Brown & Sinclair, *supra* note 65 at ch 3.

113. For further critical reflection on these methodological issues in the field of political philosophy, see David Miller, "Political philosophy for Earthlings" in David Leopold & Marc Stears, eds, *Political Theory: Methods and Approaches* (Oxford University Press, 2008) 29.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

point of applied normative philosophy is to provide sound philosophical principles for making good policy as opposed to furnishing rationalisations or 'spin' for decisions that have effectively already been taken on some other basis. Of course, political philosophy also has something to say about democratic values, and so if a policy decision is taken on grounds that the government ought to be responsive to what voters want, then there could be a democratic principle at play. But if that is the case, then government should come clean about which type of normative principle is really doing the heavy lifting and not pretend that the free speech principle is driving the policy decision when in reality it is not.

The final implication I want to tease out concerns compromise. At first glance, a philosophical free speech principle does not look like something that can or should be compromised. So, for example, if a free speech principle tells us that a certain justification for limiting speech is illegitimate and the principle is true, then that is all there is to say. But what if there is not one solitary free speech principle but a slew of free speech principles, each invoking a different fundamental value, for example, and what if two or more such principles are mutually conflicting?

The answer here might seem obvious to some people. A rational agent should use reason, argument, and evidence to determine which is the true free speech principle or which iterations are members of the set of true free speech principles, and which are members of the set of false principles. By a process of elimination, one could in theory drastically reduce the number of true or epistemically credible free speech principles, perhaps starting by examining any mutually conflicting principles. However, the main obstacle to this eliminative approach is reasonable disagreement.

Consider three examples of such disagreement. The first example concerns distinctiveness. Having looked at the available evidence and arguments, some free speech scholars are sceptical about the truth of any free speech principle that says speech protections are special partly in virtue of the fact that they can be justified in a distinctive way, such as by a morally relevant feature that does not justify (in the same way or with the same force) other types of freedoms or liberty in general.[114] Then again, not everyone shares this scepticism. Having looked at similar evidence and arguments, other free speech scholars believe that the distinctiveness of speech protections and how they are justified can indeed be demonstrated, such that there are valid reasons why free speech deserves special protection over and above the protection afforded to other freedoms.[115] Thus, the issue of whether free speech should be considered, philosophically speaking, a special right remains a matter of reasonable disagreement, and

114. See Schauer, *supra* note 2; Schauer, *supra* note 60; Alexander, *supra* note 2; Lawrence Alexander, "Is Freedom of Expression a Universal Right?" (2013) 50:3 San Diego L Rev 707; Alexander & Horton, *supra* note 11.

115. See Greenawalt, *supra* note 26; Kendrick, *supra* note 25.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

this is partly due to reasonable disagreement about the meaning of the concepts of distinctiveness, justification, and a special right.

To give a second example, consider reasonable disagreement concerning whether the true free speech principle is the (negative) extending principle that says the enforcement of civility norms is *not* a fitting justification for restricting speech, even hate speech,[116] or is instead the (positive) limiting principle that says the enforcement of civility norms *is* a fitting justification for restricting speech, especially hate speech.[117]

As a third example, even if scholars could agree on a limiting free speech principle that says, for example, if there is minimally adequate evidence that hate speech is sufficiently harmful, if restrictions on hate speech are effective and the least restrictive alternative, and if there are negligible unintended consequences, then restrictions on hate speech can be warranted, there remain grounds for reasonable disagreement concerning whether these empirical conditions have been met for any given restriction.[118] This reasonable disagreement might in the end turn on the meaning of 'minimally adequate evidence', 'sufficient harm', 'effective', 'restrictive alternatives', and 'negligible unintended consequences'.

I believe the implication is clear: As a society (including our lawmakers and judges but also our philosophers) we may sometimes have to reach compromises over our free speech principles. By this I mean not merely compromise for the sake of other sorts of normative principles (i.e., non-free speech principles like the dirty hands principle) or even for considerations that are not normative principles at all; I also mean compromise for the sake of bringing the set of free speech principles into harmony with each other. Talk of a person 'compromising his or her principles', or public institutions 'compromising relevant principles' can have a pejorative ring. That compromise may have a bad reputation in some fields of law and politics, and this might have something to do with how it is sometimes characterised. Consider Eric Heinze's negative portrayal: "Processes of compromise have long been acknowledged within legislatures and political parties, the natural homes of winks and nods, handshakes and

116. See Post, *supra* note 39.
117. See Rice, *supra* note 11.
118. For the view that these empirical conditions can be, and often are, met, although not always and not necessarily easily, see e.g. Brown, *supra* note 46 at chs 3-9; Katharine Gelber & Luke McNamara, "Evidencing the harms of hate speech" (2016) 22:3 Social Identities 324; Katharine Gelber, "Hate Speech—Definitions & Empirical Evidence" (2017) 32:3 Const Commentary 619; Brown, *supra* note 57; Alexander Brown & Adriana Sinclair, *Frontiers of Hate Speech: Exploring the Limits of the Ordinary and Legal Concepts* (Cambridge University Press, 2024) at 514, n 35. For a high level of scepticism over whether these empirical conditions tend to be met in the real world, see Richard L Abel, *Speaking Respect, Respecting Speech* (University of Chicago Press, 1998) at ch 6; Baker, *supra* note 24; Heinze, *supra* note 23; Weinstein, *supra* note 24; James Weinstein, "Viewpoint Discrimination, Hate Speech, and Political Legitimacy: A Reply" (2017) 32:3 Const Commentary 715; Nadine Strossen, *Hate: Why We Should Resist It with Free Speech, Not Censorship* (Oxford University Press, 2018); Tara Beattie & Gavin Phillipson, "Criminalising Pornography" in Ian Loveland, ed, *British and Canadian Public Law in Comparative Perspective* (Bloomsbury, 2021) ch 6.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

horse-trading."[119] However, I think there are two good reasons to see compromise as something other than suspect or grubby.

One reason has to do with the nature of principles. Talk of compromising over free speech principles might not be such a bad thing if one understands that some principles, even philosophical free speech principles, are different from certain kinds of rules. Some rules operate as side-constraints, such that if the rule is valid then it must be observed. A coherent system of such rules requires that they be mutually consistent, meaning that agents can observe all the rules at the same time because the decisions required by the rules can be jointly performed. Some principles, by contrast, might have differing degrees of weight or importance within a system. Ronald Dworkin argues that legal principles have this quality—the quality of having degrees of weight or importance—and it seems not unreasonable to think some philosophical principles also have it.[120] Thus, it could be coherent to speak in terms of 'compromising' less weighty or important normative principles for the sake of more weighty or important principles. Interestingly, some philosophical free speech principles may share this quality in common with some doctrinal free speech principles. Consider the fact that chapter 5 of Mill's *On Liberty* appears to contain exceptions to his own harm principle (e.g., public acts of indecency)[121] and the fact that the First Amendment doctrine of viewpoint and content neutrality also has exceptions,[122] suggesting they are not side-constraints after all but instead principles that have a degree of weight or importance. (Of course, the exceptions could also suggest these are rules but more complex rules than they first appear. Indeed, when it comes to philosophical free speech principles, Dworkin sometimes speaks as though the principles he defends are rules (side-constraints) or as though, if they are principles, then they are principles with absolute weight or importance, such that they take lexical priority over all other free speech principles.[123] But let us put the worry about complex rules disguised as principles aside and just assume that at least some landmark First Amendment cases can be usefully interpreted as involving compromise over doctrinal free speech principles.)[124]

119. Eric A Heinze, Book Review of *Hate Speech Law: A Philosophical Examination* by Alexander Brown, (2018) 8 International Dialogue 85 at 86.

120. See Ronald M Dworkin, "The Model of Rules" (1967) 35:1 U Chicago L Rev 14 at 27.

121. For further discussion of these putative exceptions, see Jonathan Wolff, "Mill, Indecency and the Liberty Principle" (1998) 10:1 Utilitas 1; Jonathan Riley, *Routledge Philosophy Guidebook to Mill On Liberty* (Routledge, 1998); Alexander Brown, "J.S. Mill & Violations of Good Manners", *Philosophy Now* 76 (1 November 2009) 12; Piers Norris Turner, "'Harm' and Mill's Harm Principle" (2014) 124:2 Ethics 299; Clare McGlynn & Ian Ward, "Would John Stuart Mill Have Regulated Pornography?" (2014) 41:4 JL & Soc'y 500.

122. See e.g. *RAV v St Paul*, 505 US 377 (1992) [*RAV*]; *Virginia v Black*, 538 US 343 (2003) [*Black*].

123. See Dworkin, "Foreword", *supra* note 24; Dworkin, "Reply to Jeremy Waldron", *supra* note 24.

124. For an interpretation of *RAV* and *Black* as involving compromises over the relevant doctrinal free speech principles, see Brown, *supra* note 46 at ch 9.5.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

Sometimes, bringing principles into harmony could be assisted by finding a singular benchmark or standard for assessing weight or importance, such that it is possible to determine the weight or importance of every principle by the same standard. And, as Heinze puts it, "such rhetoric of 'compromise' seems merely to mask age-old processes of balancing incommensurables."[125] Moreover, this approach relies on finding a benchmark or standard that is up to the task. One option might be to ask questions such as 'What is the underlying point and purpose of free speech principles?' and 'Which principles better serve that point and purpose?'. Yet there is still likely to be reasonable disagreement here—reasonable disagreement about what the relevant purpose should be and about what it means to serve it well, better, or best. As a result, it could be difficult to say with any substantial level of reasonable consensus which principles should be compromised for which other principles. In addition, even if there were an obvious benchmark by which to measure weight or importance, it could be that for some pairwise comparisons of free speech principles the right thing to say is simply that they have the same or equal weight or importance. In these eventualities, talk of compromising free speech principles can no longer be a rhetorical allusion to balancing; it really means what it says, compromise.

A second reason why talk of compromise over free speech principle might not be such a bad thing is if the compromise is done in the 'right' way. This is not a point about providing a singular benchmark or standard for assessing weight or importance. It is a much broader point about the entire process of reaching compromise; about how the art of compromise is conducted (by actual people). Even if resolving policy dilemmas over free speech policies on a principled basis may require some degree of compromise over mutually-conflicting free speech principles, such compromise can be much more than merely "winks and nods, handshakes and horse-trading"; it can be compromise that lives up to certain ideals of compromise. Arguably, what is needed here is principled compromise in the twin sense of compromise over free speech principles and compromise that is principled, as in, compromise that is conducted in accordance with appropriate meta-level principles, such as principles of deliberative democracy, for instance. There has already been research into this sort of principled compromise, and clearly more is needed.[126] But a common thread is the meta-level principle that within the process of reaching compromises over principles, the agents concerned must give equal space to each other to propose and defend principles (equality), must defend principles based on mutually acceptable reasons (reciprocity), must enter the compromise process in the right spirit (respect), and must be willing to negotiate over what the standards of equality, reciprocity, and respect should be (reflexivity). If there is a more important

125. Heinze, *supra* note 119 at 91.

126. See Brown, *supra* note 46 at ch 10; Daniel Weinstock, "Compromise, pluralism, and deliberation" (2017) 20:5 Critical Rev Intl Social & Political Philosophy 636.

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press

meta-level principle (a principle about reaching compromises over free speech principles), then I do not know what it is.

Acknowledgments: An early version of this article was presented at a workshop on free speech organised by Robert Simpson at the Department of Philosophy, University College London, in June of 2019. I would especially like to thank Christopher Cousens, Jeffrey Howard, Maxime Lepoutre, Gavin Phillipson, and Robert Simpson for comments and suggestions.

Alexander Brown is Associate Professor of Political and Legal Theory at the University of East Anglia (UEA). He is the author of *Hate Speech Frontiers: Exploring the Limits of the Ordinary and Legal Concepts* (Cambridge University Press, 2023); *An Ethics of Political Communication* (Routledge, 2021); *The Politics of Hate Speech Laws* (Routledge, 2020); *A Theory of Legitimate Expectations for Public Administration* (Oxford University Press, 2017); *Hate Speech Law: A Philosophical Examination* (Routledge, 2015); *Personal Responsibility: Why it Matters* (Continuum, 2009); and *Ronald Dworkin's Theory of Equality: Domestic and Global Perspectives* (Palgrave, 2009). Email: alexander.c.brown@uea.ac.uk

https://doi.org/10.1017/cjlj.2024.20 Published online by Cambridge University Press