

# After Christchurch: Hate, harm and the limits of censorship

## 6. Striking a fair balance when regulating harmful communication

*David Bromell*

Working Paper 21/07



Institute for Governance  
and Policy Studies  
A research institute of the School of Government



INSTITUTE FOR GOVERNANCE AND  
POLICY STUDIES  
WORKING PAPER  
21/07

MONTH/YEAR

April 2021

AUTHOR

David Bromell  
Senior Associate  
Institute for Governance and Policy Studies

INSTITUTE FOR GOVERNANCE AND  
POLICY STUDIES

School of Government  
Victoria University of Wellington  
PO Box 600  
Wellington 6140  
New Zealand

ACKNOWLEDGEMENT

For any queries relating to this working paper,  
Research on this series of working papers has  
been financially supported by a fellowship at the  
Center for Advanced Internet Studies (CAIS) in  
Bochum, NRW, Germany (Oct 2020—Mar 2021).

DISCLAIMER

The views, opinions, findings, and conclusions or  
recommendations expressed in this paper are  
strictly those of the author. They do not  
necessarily reflect the views of the Institute for  
Governance and Policy Studies, the School of  
Government, Victoria University of Wellington, or  
the Center for Advanced Internet Studies (CAIS).

This is paper six in a series of seven working papers, **After Christchurch: Hate, harm and the limits of censorship**.

The series aims to stimulate debate among policy advisors, legislators and the public as New Zealand considers regulatory responses to ‘hate speech’ and terrorist and violent extremist content online following the terrorist attack on Christchurch mosques in March 2019 and the Royal Commission of Inquiry that reported in November 2020.

The seven working papers in this series are:

Title	Reference
<b>1. The terrorist attack on Christchurch mosques and the Christchurch Call</b>	WP 21/02
<b>2. ‘Hate speech’: Defining the problem and some key terms</b>	WP 21/03
<b>3. Challenges in regulating online content</b>	WP 21/04
<b>4. Regulating harmful communication: Current legal frameworks</b>	WP 21/05
<b>5. Arguments for and against restricting freedom of expression</b>	WP 21/06
<b>6. Striking a fair balance when regulating harmful communication</b>	WP 21/07
<b>7. Counter-speech and civility as everyone’s responsibility</b>	WP 21/08

From October 2020 to March 2021, Dr David Bromell was a research Fellow at the Center for Advanced Internet Studies (CAIS) in Bochum, North Rhine-Westphalia, Germany, which supported his research on this series of working papers. He is a Senior Associate of the Institute for Governance and Policy Studies in the School of Government at Victoria University of Wellington, and a Senior Adjunct Fellow in the Department of Political Science and International Relations at the University of Canterbury. From 2003 to 2020 he worked in senior policy analysis and advice roles in central and local government.

He has published two monographs in Springer’s professional book series:

- *The art and craft of policy advising: A practical guide* (2017)
- *Ethical competencies for public leadership: Pluralist democratic politics in practice* (2019).

## Contents

Abstract.....	5
Introduction .....	5
Public and private .....	7
Harm and offence .....	9
Persons and groups.....	10
Power, numbers and the heckler’s veto .....	11
Neither neutrality nor censorship.....	13
Learn from what does and does not work.....	14
Conclusion.....	15
References .....	15

## Striking a fair balance when regulating harmful communication

### Abstract

This paper encourages policy makers to strike a fair balance when regulating harmful communication. It discusses some important distinctions between public and private communication, harm and offence, and persons and groups.

Previous working papers in this series summarise the commitments made by governments and online service providers in the Christchurch Call (Working paper 21/02), define the problem of ‘hate speech’ and other key terms (Working paper 21/03), identify challenges in regulating online content (Working paper 21/04), summarise current legal frameworks for regulating harmful communication (Working paper 21/05), and analyse arguments for and against restricting freedom of expression (Working paper 21/06).

Harmful communication is public communication that incites discrimination, hostility or violence against members of a social group with a common ‘protected characteristic’ such as nationality, race or religion. This is the proper focus of regulation. In many jurisdictions, it can be and is dealt with under existing criminal law.

Any proposal to regulate private communication, or to protect social groups and their beliefs, values and practices from criticism and offence, exacts too high a price because of the extent to which this restricts freedom of expression. Policy makers need to refrain from using the coercive power of the state to enforce a ‘heckler’s veto’—or a ‘mourner’s veto.’ In a free, open and democratic society, agonistic respect and toleration, including and especially toleration of ideas, beliefs, attitudes and practices which we dislike, are preferable to ‘calling out’, ‘cancel culture’ and de-platforming.

This does not necessarily imply state neutrality. Governments can deploy expressive rather than coercive powers to address ‘lawful hate speech’, upholding the ideal of free and equal citizenship and learning from what works and does not work to maintain public order and promote civility.

This anticipates the final paper in this series, on counter-speech strategies as alternatives, or complements, to prohibition and censorship, and civility as everyone’s responsibility (Working paper 21/08).

**Tags:** #ChristchurchCall #censorship #hatespeech #freespeech #freedomofexpression

### Introduction

Working paper 21/06 defined public policy making in a free, open and democratic society as incremental social problem solving as we work out the right thing to do, or not do, through agreed processes. It demands the prudential balancing of divergent, competing and conflicting interests, exercising principled pragmatism in local contexts. Writing about freedom of speech, Dieter Grimm (2009, p. 13) comments that ‘in the final analysis, everything depends on a fair balance between the rights and interests at stake’.

When policy makers consider options to regulate harmful communication, important distinctions are between public and private communication, harm and offence, and persons and groups. This in turn

demands clear thinking about the differences between disagreement and discrimination, criticism and incitement, satire and provocation—and between portrayal, promotion and enactment (Edgar, 2009, p. 585). As Nadine Strossen (2018, p. 67) argues:

Government may not regulate ‘hate speech’ solely because its message is disfavored, disturbing, or feared. But government may restrict ‘hate speech’ when it directly, demonstrably, and imminently causes certain specific, objectively ascertainable serious harms.

The Federation of Islamic Associations of New Zealand, in its submission to the Royal Commission of Inquiry on the terrorist attack on Christchurch mosques, commented:

... there is a clear line between the freedom to openly and publicly discuss, defend, and teach different positions on moral and social issues and theological matters, and professing hatred and inciting violence. While the former are necessary for the maintenance and progress of a free and democratic society, the latter are destructive for a free and democratic society (FIANZ, 2020, p. 125).

Drawing a clear line between permissible and impermissible communication is not, however, a simple or straightforward matter. As noted in Working paper 21/04, **Challenges in regulating online content** (p. 8), one response to clamping down on harmful communication is to cloak ideology in generalities and to use coded and covert signals, irony, sarcasm, images, memes and in-jokes to shroud borderline content and extreme views in plausible deniability. This allows users to operate right up to the limits of platform rules and guidelines. It positions the alt-right, for example, as a counterculture and depicts anyone who is offended as ‘a triggered, liberal snowflake who doesn’t get the joke’ (Miller-Idriss, 2020a, p. 66).

The difficulty of clearly delineating harmful communication means that regulation is always and especially a balancing act (Gärditz, 2020)—between freedom of expression, protection from harm, promotion of social cohesion, maintenance of public order and ensuring that the law can practically be enforced (Royal Commission of Inquiry, 2020, p. 701).

The People’s Republic of China provides an object lesson in how not to strike a fair balance when regulating freedom of expression. Daniela Stockmann (2020) comments:

China’s highly sophisticated tools for policing content, blocking websites and tracking user behaviour have earned its policies notoriety as The Great Firewall of China. China has built an extensive institutional and technological infrastructure to monitor, surveil, delete and block content by users. This is no blueprint for democracies grappling with disinformation, hate speech and election meddling.

Regulation of harmful communication in a free, open and democratic society requires *proportionality*—the least coercive measures possible, and as a last, not first, resort of government for preventing harm (Sumner, 2009, p. 207).

## Public and private

While digitisation blurs boundaries between ‘public’ and ‘private’ space (Working paper 21/04, pp. 7–8), regulation of harmful communication needs to maintain a distinction between private and public communication.

The UK Public Order Act 1986 s18, for example, specifies that:

(2) An offence under this section may be committed in a public or a private place, except that no offence is committed where the words or behaviour are used, or the written material is displayed, by a person inside a dwelling and are not heard or seen except by other persons in that or another dwelling.

(4) In proceedings for an offence under this section it is a defence for the accused to prove that he was inside a dwelling and had no reason to believe that the words or behaviour used, or the written material displayed, would be heard or seen by a person outside that or any other dwelling (UK Government, 2006).

Something is ‘public’ if it directly or indirectly concerns, or could potentially concern, any member or members of a community indiscriminately (Barry, 1962, pp. 195–196; Bromell, 2017, p. 59) and is ‘open to witness’:

The public is the space in which witnessing can take place. Conversely, one is a private being—a solely personal actor—when one’s actions cannot be witnessed by others. The private sphere is the domain in which one can only be witnessed by intimate observers (Coleman & Ross, 2010, p. 5).<sup>1</sup>

We can further distinguish public from private in terms of the direct and indirect consequences of actions. This is critical when assessing whether government intervention is justifiable to prevent or respond to something that incites, or is likely to incite, disturbance of public order that causes harm.<sup>2</sup> Barry (1965, pp. 191–192), following Bentham, distinguished private, reflective, semi-public and public offences (or injuries).

- A *private injury* damages one or more identifiable individuals.
- A *reflective injury* damages one’s own self.
- A *semi-public injury* affects a portion of the community (a ‘public’) and, depending on the duration and severity of the offence, may justify government action.
- A *public injury* produces some actual or potential danger either to all members of a state, or to an indefinite number of non-assignable individuals (anyone at all) in a specific context who may be affected by the consequences of an action.

---

<sup>1</sup> See also Asen (2000). Thomas Nagel laments a decline of respect for the boundaries between the private and the public, concealment (or at least reticence and privacy) and exposure: ‘The liberal idea, in society and culture as in politics, is that no more should be subjected to the demands of public response than is necessary for the requirements of collective life’ (Nagel, 2002, p. 13).

<sup>2</sup> Cf. New Zealand’s Films, Videos, and Publications Classification Act 1993 s3, which defines a publication as ‘objectionable’ if it ‘describes, depicts, expresses, or otherwise deals with matters such as sex, horror, crime, cruelty, or violence in such a manner that the availability of the publication is likely to be injurious to the public good’.

Harmful communication is public communication that is, or is intended or likely to become, open to witness. It intends or is likely to incite discrimination, hostility or violence against assignable or non-assignable individuals (semi-public or public injury), based on some characteristic shared by or attributed to members of some actual or supposed social group with a common 'protected characteristic' such as nationality, race or religion.

Respect for distinctions between private and public communication needs to be balanced, however, even within the family, workplaces and voluntary associations, by what Corey Brettschneider (2012, pp. 26–27) calls a principle of 'public relevance':

What makes certain beliefs and practices publicly relevant is that they conflict with our public status as free and equal citizens. According to the principle of public relevance, these beliefs and practices are of public concern, and ideally should be changed to make them consistent with the ideal of free and equal citizenship ... Although the family and civil society should be protected by rights against coercion, their beliefs and practices should be open to criticism if they are hostile to free and equal citizenship.

Brettschneider argues, for example, that 'raising children as racists does not justify coercive intervention within the family or the forcible removal of the child from the home, but it is publicly relevant and may call for a persuasive response by fellow citizens and the state' (p. 28).

Carolyn Evans (2009, p. 363) similarly reflects that while religious speech that undermines gender equality often occurs in the private realm of church, temple, synagogue, mosque or home, this is where some women live a great deal of their lives: 'At the extreme, women in some small, close-knit, conservative religious communities may have almost no life in what the liberal world-view conceives of as the public realm.'

Evans welcomes transgressive, transforming speech and women's movements within religions, and proposes reforms that are generally non-coercive and respectful of religious freedom, while still seeking to influence religious groups to take gender equality seriously by, for example:

- Providing high-quality education that creates conditions in which women are given the tools and opportunities to make a real choice about religious issues;
- Providing strong protection against domestic violence, with no exceptions for religious or cultural justifications for the perpetuation of such violence;
- Permitting tax exemptions only to religious groups that voluntarily accept compliance with discrimination laws; and
- Requiring religious institutions to demonstrate that they have engaged in a consultative process that includes women before accepting any submissions on public policy as representative of the views of their community.

There is a cautionary note to be sounded here for religious groups that urge government to criminalise 'hate speech'—the legislative measures they demand to protect them also apply when, for example, they advocate discrimination against people based on their sexual orientation (Leigh, 2009).



## Harm and offence

Kwame Anthony Appiah (2012, p. 169) notes two different threats to public order from the use of language calculated to offend:

One sort of offensive speech consists of the sort of insult that might lead a normal person exposed to it to be so outraged that she would respond violently to her insulters. From this point of view, it is the consequences for public order—or rather disorder—that flow from such speech that matters, not the protection of the victim from verbal abuse. A rather different idea is that words can have the aim or effect of inflaming passions in ways that will lead to crimes against people or property, especially crimes against those defamed. Here, protection of the defamed from physical harm is at the center of concern, and the speech amounts to incitement to violence against them.

The second sort of harm, ‘stirring up’ or incitement of discrimination, hostility or violence, is the proper focus of legislation against harmful communication. In many jurisdictions, it can be and already is dealt with under existing criminal law.

The first sort of harm poses more difficult questions for policy makers, particularly when terrorist acts are performed in reaction to offence (Bromell & Shanks, 2021). A recent example is the beheading of French teacher Samuel Paty on October 16, 2020 after he had discussed with his class caricatures of Islam’s Prophet Muhammad. The killer sent a Twitter post with a picture of the severed head addressed to President Emmanuel Macron, ‘leader of the infidels’, and boasted of killing ‘one of your hell dogs who dared to denigrate Mohammed’. The 18-year-old killer, a Chechen immigrant, acted alone but was connected to others (Nossiter, 2020b; Onishi, Méheut & Foroudi, 2020). He appears to have been motivated by content circulating on social media to travel to Conflans-Saint-Honorine to kill Paty. French government spokesperson Gabriel Attal said: ‘Things began on social media and they ended on social media’ (Mallet & Murphy, 2020).

This was the second terrorism-related incident since Charlie Hebdo, the French satirical newspaper, published a symbolic cover with Muhammed cartoons to mark the start of the trial of 14 alleged accomplices in the January 2015 terrorist attack on the newspaper, in which twelve people, including nine journalists, were killed (Ganley, 2020; McAuley, 2020; Nossiter, 2020a). And that was not the end of it, with subsequent Islamist terror attacks in Nice on October 29 and in Vienna on November 2, 2020.

Criticism and satire are not ‘harmful communication’ but can indirectly cause harm when offence prompts incitement to retaliatory acts of terrorism. This is the difficult space within which policy makers need to strike a balance between freedom of expression and prevention of harm.

Holmes (2012, p. 349) reflects that ‘one of the purposes of criminal law is to control the revenge instinct in society.’ But in weighing up the choices involved, Appiah (2012, pp. 169–170) comments:

When we have to decide between regulating expression, on the one hand, and allowing it while controlling the violence it may elicit, on the other, it matters both what the costs of each will be and whether we think the violence of the responses is reasonable ... A concern for public order can raise the issue whether speech should be barred; but even if speech poses a threat to public order, it may be worth expending the resources to allow it. If freedom of speech is worth something, sometimes the public should pay the costs of allowing it.

Bret Stephens (2020), writing in the *New York Times*, cautions against defaulting to a middle position that goes like this:

Fanatics shouldn't kill people, and writers and artists shouldn't needlessly offend fanatics. It's a compromise that is fatal to liberalism. It reintroduces a concept of blasphemy into the liberal social order. It gives the prospectively insulted a de facto veto over what other people might say. It accustoms the public to an ever-narrower range of permissible speech and acceptable thought.

Regulation to protect people from offence (that is, psychological harm) exacts an extremely high social cost:

A world in which we were constantly trying to predict and avoid upset would be a world with much less communication, and many of the desirable consequences of a world rich in communication would have to be foregone. Identifying actual psychological harms is, in any case, a difficult exercise, one that we would be wise to avoid imposing on our courts. And making any upset caused by speech or writing a legally cognizable harm incentivizes people to pretend to be upset to shut others up (Appiah, 2012, p. 171).

New Zealand political commentator Chris Trotter (2020) has put it bluntly:

Attempting to pass laws against the giving of offence ... is a fool's errand. Far from eliminating offensiveness, such laws will only encourage and intensify it. Harm cannot be prevented, but it can be healed. Building trust and amity between peoples is achieved by starting conversations—not by shutting them down.

## Persons and groups

Although he analyses 'hate speech' as a form of group defamation, Jeremy Waldron (2012, p. 60) insists that regulation of 'hate speech' is to protect the dignity of the individual person 'when defamatory imputations are associated with shared characteristics such as race, ethnicity, religion, gender, sexuality, and national origin', not the dignity of a culture, religion or other social group.

In other words, when regulating harmful communication, it is important to distinguish between, on the one hand, regulation to protect persons as members of social groups from harmful communication that incites discrimination, hostility or violence against them and, on the other hand, regulation designed to immunise a social group and its beliefs, values and practices from criticism and offence.

Dieter Grimm reminds us that the secular, liberal state was a reaction to the devastating wars of religion that followed the Reformation of the sixteenth century. Separating religion and the state provides religious freedom, limited only by the requirement that religions, which often have mutually exclusive beliefs, co-exist peacefully. Any such limitation must be determined by law within a democratic process, which in turn requires an unimpeded public discourse:

This means that, in the secular state, religion cannot be exempted from criticism. Freedom of speech and freedom of the media is not less important than freedom of religion. The more a religious group claims public recognition and respect for its religious norms, the more it must be exposed to public discourse. Consequently, there are some limits to protection of religious beliefs against speech (Grimm, 2009, p. 18).

Kenan Malik (in Molnar, 2012, p. 82) has argued strenuously against religious defamation laws:

Proponents of religious defamation laws suggest that religion is not just a set of beliefs but an identity, and an exceptionally deeply felt one at that. It is true that religions often form deep-seated identities. But then, so do many other beliefs. Communists were often wedded to their ideas even unto death. Many racists have an almost visceral attachment to their beliefs. Should I indulge them because their views are so deeply held? And while I do not see my humanism as an Identity with a capitol 'I,' I would challenge any Christian or Muslim to demonstrate that my beliefs are less deeply held than theirs.<sup>3</sup>

Malik further argues that demands for protection of faith or culture often come from conservatives who want to suppress dissent within their own culture or community of faith:

What is often called offense to a community or a faith is actually a debate within that community or faith. In accepting that certain things cannot be said because they are offensive or hateful, those who wish to restrict free speech are simply siding with one side in such debates—and usually the more conservative, reactionary side (p. 83).

Further, if protection from harmful communication is extended to social groups, including religious faiths, then claims to protection will almost certainly extend to nations and governments. Pro-Israel advocacy groups have, for example, called for Facebook to 'fully adopt' the International Holocaust Remembrance Alliance's (IHRA) working definition of antisemitism and to use this to exclude criticism of Israel from the bounds of acceptable discourse (Friedman, 2020). The American Civil Liberties Union has noted with concern 'attempts to suppress the speech of people on only one side of the Israel-Palestine debate' (Waheed & Hauss, 2018).

In the event, Facebook did change its policy on 'hate speech' in ways it says take the IHRA definition into account, and it has recently banned denial of the Holocaust (Bickert, 2020; Strack, 2020), but Facebook has resisted adopting the IHRA definition in its entirety. This has disappointed IHRA-focused advocates, who want Facebook not only to de-platform antisemitism, but to de-platform criticism of Israel (Friedman, 2020).

## Power, numbers and the heckler's veto

In striking a fair balance between protection from harm and freedom of expression, policy makers need to take account of how power is exercised within a society between majority and minority social groups. Policies and laws should not be determined solely by the numbers, or by who can shout (or 'mourn') the loudest.<sup>4</sup>

As noted in arguments for restricting freedom of expression in Working Paper 21/06 (p. 11), while there is a lack of empirical evidence for this, we might provisionally assume that harmful communication tends to silence the 'voice' of those targeted and to discourage them from active participation in and contribution to democratic deliberation and public life.

---

<sup>3</sup> Malik (in Molnar, 2012, p. 83) adds: '... the attempt to restrict defamation of religion is, ironically, an attack not just on freedom of speech but on freedom of worship too—and not least because one religion necessarily defames another. Islam denies the divinity of Christ; Christianity refuses to accept the Qur'an as the word of God. Each Holy Book blasphemes against the others.' See further Adams (2012).

<sup>4</sup> On the 'mourner's veto', see Ferguson (2020) and the discussion in Working paper 21/03, '**Hate speech**': **Defining the problem and some key terms**, p. 14. Ferguson provides some useful suggestions for responding to people who make emotional or emotive attempts to suppress controversial or unpopular speech.

The difficulty is that harmful communication that incites discrimination, hostility or violence also often comes from within minority social groups and the fringes of 'mainstream' society. In a free, open and democratic society, they too have a 'voice' and should not be excluded from active participation in and contribution to democratic deliberation and public life.

There is a regrettable tendency across the political spectrum to think, speak and behave as if social problems can be solved by an 'us' wishing a 'them' into silence, invisibility, powerlessness or non-existence: 'They (whoever they are) are not us.' This 'get-rid-of' mentality is just another assault on human dignity and a bullying abuse of power.

Regardless of whether individuals affiliate with the radical left or the radical right, or how repugnant others may find their beliefs, values and practices, we are all fellow citizens, 'born equal in dignity and rights' (Universal Declaration of Human Rights, Article 1). In a liberal democracy, we don't all need to like each other. We just need to respect one another's fundamental human equality and learn to live together in all our difference, resolving our conflicts politically without resorting to domination, humiliation, cruelty or violence.

Chris Trotter is one of a number of New Zealanders who formed the Free Speech Coalition, initially to oppose the de-platforming by Auckland Mayor Phil Goff of two Canadian speakers, Lauren Southern and Stefan Molyneux, from council-owned venues due to their political views (Free Speech Coalition, n.d.). The Coalition has since turned its attention to calls for censorship and new 'hate speech' laws in the wake of the Christchurch terror attacks. Trotter (2020) has reflected that:

It is one of the greatest tragedies of contemporary 'left-wing' politics: that its practitioners have allowed themselves to become identified, irretrievably, with the suppression of free speech. Most particularly, with the suppression of the free speech of persons identified as 'right wing', or, more ludicrously, as 'Nazis' and 'fascists'. Worse still, they have secured this 'de-platforming' by threatening to unleash violence and disorder if these individuals are permitted to speak. They have thus supplied local government, university and corporate leaders with the 'health and safety' justification for shutting these speakers down. Free speech advocates refer to this tactic as 'The Thug's Veto'.<sup>5</sup>

Social media has made it all too easy to name-call and abuse people within digital 'echo chambers'—labelling them as 'racist' or '-phobic', engaging in *ad hominem* attacks and belittling them. Carolin Emcke (2019) cautions against meeting hatred with hatred. A free, open and democratic society can defend itself, without defeating itself, only by remaining liberal and open, and by addressing the social and economic conditions that generate hatred and violence:

Because hatred can only be contested with what the haters lack: careful observations, unstinting precise distinctions, and self-doubt. That means slowly breaking hatred down into its component parts, separating the momentary feeling of hatred from its ideological prerequisites, and observing how it arises and operates in a specific historical, regional, cultural context (p. xvi).

---

<sup>5</sup> A 'thug's veto' (or 'heckler's veto') is when activists force the cancellation or disruption of an event ('cancel culture') or the silencing of a speaker ('de-platforming') they don't agree with. De-platforming in this sense means something different from the technical act of disabling social media profiles and blocking or removing content as discussed in Working paper 21/04, **Challenges in regulating online content**, pp. 16–23). See Strossen (2018, pp. 18–19) and Working paper 21/05, **Regulating harmful communication: Current legal frameworks**, p. 7, fn.1. For an analysis of an incident of university de-platforming in New Zealand, see Bromell, 2019, pp. 84–88.

For Emcke, this includes respecting pluralism and individuality, the public and the private:

*Really existing in the plural means mutually respecting everyone's individuality and uniqueness. I do not have to live or think exactly as everyone else does. I do not have to share the practices and beliefs of others. I do not have to like them or understand them. This too is an aspect of the enormous freedom of a truly open, liberal society: not having to like one another, and yet being able to let one another be (Emcke, 2019, p. 114, emphasis hers).*

This suggests strategies of 'agonistic respect', including creating political space even for extreme views (M. Malik, 2009, pp. 109–111) under the rule of law and with a view to *engagement* that opens up internal and external critique of extremist ideologies.

The word *agonistic* originally related to athletic contests in ancient Greece. It has a dictionary sense of being combative, and of striving to overcome an opponent in argument. Agonistic pluralism as a political theory accepts that conflict is a permanent feature of politics. It pays particular attention to *hegemony*, the dominance of one social group over another.

Democratic debate is a process of struggle, in which citizens engage in contest with one another. Iris Young encouraged us to work our conflicts out in local contexts and negotiate priorities and trade-offs through an exchange of public reasons. This implies a politics of difference without exclusion; but equally it implies a politics without community, a politics of unassimilated otherness, a togetherness of strangers (Young, 1990, p. 237).

Chantal Mouffe contrasts *agonism* with *antagonism*. She also accepts that conflict and division are inherent to politics: '... the specificity of modern pluralist democracy—even a well-ordered one—does not reside in the absence of domination and violence but in the establishment of a set of institutions through which they can be limited and contested' (Mouffe, 2005a, p. 22).<sup>6</sup>

## Neither neutrality nor censorship

Corey Brettschneider (2012) argues for neither neutrality nor prohibition (state censorship) in relation to discriminatory or hateful viewpoints. He argues for a third way (democratic persuasion) that distinguishes between a state's coercive power, or its ability to place legal limits on harmful communication, and its expressive power, or its ability to influence beliefs and behaviour: 'On my view, the state should simultaneously protect hateful viewpoints in its coercive capacity and criticize them in its expressive capacity' (p. 3):

Unlike militant democrats, I do not support coercive restrictions on hate speech, provided there is no direct threat or danger of imminent harm. But I do suggest that hateful or discriminatory views violate the democratic ideal and should be publicly criticized (pp. 63–64).

Brettschneider argues (p. 89) that democratic persuasion expressed by the state needs to be subject not only to a means-based limit but also to a substance-based limit:

It is ... essential to the means-based limit that even when the state is using democratic persuasion to criticize hateful viewpoints, it does not undermine the status of citizens who

---

<sup>6</sup> On agonism and a pluralist democratic politics, see further Bromell, 2019, pp.36–39; Mouffe, 2005b, 2009, 2013; Connolly, 2002; Young, 2000; and Laclau & Mouffe, 1985.

hold these viewpoints as free and equal bearers of rights. The means-based limit therefore emphasizes using reasons to attack hateful beliefs, arguments, or positions, but it should avoid demonizing individuals or exiling them from society.

The substance-based limit restricts the kind of beliefs the state is rightly concerned to transform through its expressive capacity, and the circumstances under which the state is justified in exercising that capacity. The state should use its expressive capacity to challenge only those beliefs that violate the ideal of free and equal citizenship. In particular, the state should not seek to transform all inegalitarian beliefs, but only those that challenge the ideal of free and equal citizenship.

Brettschneider (2012, p. 85) argues that this obligation to engage in democratic, non-coercive persuasion falls squarely on the state, as well as on voluntary associations, private actors (including workplaces) and individual citizens.

## Learn from what does and does not work

Cynthia Miller-Idriss, director of the polarisation and extremism research and innovation lab (PERIL) at the American University, recommends learning from what has worked and not worked in government responses to radicalisation, extremism and harmful communication. She draws a parallel, for example, with public health campaigns—messaging on eating habits and their long-term impact on health appears to have little or no impact, but campaigns that raise awareness of manipulative marketing by fast-food companies have changed behaviours, especially among boys. She reflects that no one likes to be manipulated, so work to counter radicalisation and extremism is more likely to succeed if it focuses on how manipulation, propaganda and disinformation work (Miller-Idriss, 2020b).

Miller-Idriss recommends learning from Germany's broad and comprehensive approach to combatting right-wing extremism, integrated into the mandates of a broad range of federal, state and local actors, including intelligence agencies, police, schools, theatre and arts programmes, sports clubs and teams, NGOs and so on:

What the German approach makes clear is that a national strategy focusing predominantly on surveillance and monitoring of extremists will never be enough. Successfully countering right-wing extremism requires deep integration of that work into the strengthening and protection of democracy itself ... Counterextremism work is integrated into every region and small town. There are hundreds of federally funded NGOs, state agencies, local organizations, initiatives, and projects nationwide working with schools, youth centers, and communities to provide outreach, counseling, and rehabilitation support to at-risk youth and drop-outs from far-right scenes ... Importantly, the German approach recognizes that effective interventions require meeting at-risk youth where they are (Miller-Idriss, 2020a, pp. 173, 174–175).

She concedes that the German approach has not solved the problem of right-wing extremism, which is rising in Germany as it is across Europe,<sup>7</sup> 'but as imperfect as the German approach may be, attacks like the ones in New Zealand or El Paso are made less likely by integrating counterextremism into broader democracy-building efforts' (Miller-Idriss, 2020a, p. 186).

---

<sup>7</sup> A recent report by Germany's Interior Ministry noted, for example, that verbal abuse and violence targeting politicians had increased in 2020, more often from left-wing than right-wing extremists (Knight, 2021).

Bhikhu Parekh (2012, pp. 54–55) similarly concludes that:

Law is most effective and the risk of its abuse is considerably reduced when it is part of a wider antidiscrimination and conciliation strategy, is accompanied by a campaign of public education, and is carefully drafted and directed against clearly defined forms of hate speech.

## Conclusion

When policy makers consider options to regulate harmful communication, important distinctions are between public and private communication, harm and offence, and persons and groups.

Regulation should focus on harmful public communication that incites discrimination, hostility or violence against members of a social group with a common ‘protected characteristic’ such as nationality, race or religion. In many jurisdictions, this can be and is already dealt with under existing criminal law.

Regulation of harmful communication can justifiably protect the dignity of persons as members of social groups with ‘protected characteristics’. Regulation cannot, however, justifiably seek to protect the dignity of a culture, religion or other social group. Legislation that attempts to regulate private communication, or to protect social groups and their cultures, beliefs, values and practices from criticism and offence, exacts too high a price because of the extent to which this restricts freedom of opinion and expression.

Policy makers need to refrain from using the coercive power of the state to enforce a ‘heckler’s veto’. In a free, open and democratic society, agonistic respect and toleration are preferable to ‘calling out’, ‘cancel culture’ and de-platforming.

This does not necessarily imply state neutrality. Governments can deploy expressive rather than coercive powers to address ‘lawful hate speech’ (Working paper 21/05, p. 9), upholding the ideal of free and equal citizenship and learning from what works and does not work to maintain public order and civility.

This anticipates the final paper in this series, on counter-speech strategies as alternatives, or complements, to prohibition and censorship, and civility as everyone’s responsibility (Working paper 21/08).

## References

- Adams, F. (2012). On American hate speech law. In M. Herz & P. Molnar (Eds), *The content and context of hate speech: Rethinking regulation and responses* (pp. 116–126). Cambridge: Cambridge University Press.
- Appiah, K. (2012). What’s wrong with defamation of religion? In M. Herz & P. Molnar (Eds), *The content and context of hate speech: Rethinking regulation and responses* (pp. 164–182). Cambridge: Cambridge University Press.
- Asen, R. (2000). Seeking the ‘counter’ in counterpublics. *Communication Theory*, 10(4), 424–446. <https://doi.org/10.1111/j.1468-2885.2000.tb00201.x>
- Barry, B. (1962). The use and abuse of ‘the public interest’. In C. Friedrich (Ed.), *Nomos V: The public interest* (pp. 191–204). New York: Atherton Press.
- Barry, B. (1965). *Political argument*. London: Routledge & Kegan Paul.

- Bickert, M. (2020). Removing Holocaust denial content. Facebook, October 12, 2020. Accessed October 19, 2020, from <https://about.fb.com/news/2020/10/removing-holocaust-denial-content/>
- Brettschneider, C. (2012). *When the state speaks, what should it say? How democracies can protect expression and promote equality*. Princeton, NJ: Princeton University Press.
- Bromell, D. (2017). *The art and craft of policy advising: A practical guide*. Cham, CH: Springer.
- Bromell, D. (2019). *Ethical competencies for public leadership: Pluralist democratic politics in practice*. Cham, CH: Springer.
- Bromell, D., & Shanks, D. (2021). Censored! Developing a framework for making sound decisions fast. *Policy Quarterly*, 17(1), 42–49. <https://doi.org/10.26686/pq.v17i1.6729>
- Coleman, S., & Ross, K. (2010). *The media and the public: 'Them' and 'us' in media discourse*. Chichester: Wiley-Blackwell.
- Connolly, W. (2002). *Identity, difference: Democratic negotiations of political paradox*. Minneapolis, MN: University of Minnesota Press.
- Edgar, D. (2009). Shouting fire from the nanny state to the heckler's veto: The new censorship and how to counter it. In I. Hare & J. Weinstein (Eds.), *Extreme speech and democracy* (pp. 583–597). Oxford: Oxford University Press.
- Emcke, C. (2019). *Against hate*. Cambridge: Polity Press. First published in German as *Gegen den Hass*, Frankfurt am Main: Fischer, 2016.
- Evans, C. (2009). Religious speech that undermines gender equality. In I. Hare & J. Weinstein (Eds.), *Extreme speech and democracy* (pp. 357–374). Oxford: Oxford University Press.
- Ferguson, C. (2020). Resisting the Mourner's Veto. *Quillette*, December 3, 2020. Accessed February 15, 2021, from <https://quillette.com/2020/12/03/resisting-the-mourners-veto/>
- FIANZ. (2020). FIANZ submission to the Royal Commission of Inquiry into the Attack on Christchurch Mosques. Federation of Islamic Associations of New Zealand Inc., February 2020. Accessed December 14, 2020, from <https://fianz.com/christchurch-mosques-attach/>
- Free Speech Coalition. (n.d.). About the Free Speech Coalition. Accessed October 19, 2020, from <https://www.freespeechcoalition.nz/about>
- Friedman, L. (2020). Israel-advocacy groups urge Facebook to label criticism of Israel as hate speech. *Jewish Currents*, August 19, 2020. Accessed October 19, 2020, from <https://jewishcurrents.org/israel-advocacy-groups-urge-facebook-to-label-criticism-of-israel-as-hate-speech/>
- Ganley, E. (2020). After teacher decapitated, French president calls for united stand against terrorism. *Stuff*, October 17, 2020. Accessed October 19, 2020, from <https://www.stuff.co.nz/world/europe/300134973/after-teacher-decapitated-french-president-calls-for-united-stand-against-terrorism>
- Gärditz, K. (2020). Die Grenze des Sagbaren. *Legal Tribune Online*, June 22, 2020. Accessed November 13, 2020, from <https://www.lto.de/recht/hintergruende/h/bverfg-beschluss-1-bvr-2459-19-grundrechte-meinungsfreiheit-beleidigung-grenze/>
- Grimm, D. (2009). Freedom of speech in a globalized world. In I. Hare & J. Weinstein (Eds.), *Extreme speech and democracy* (pp. 11–22). Oxford: Oxford University Press.
- Holmes, S. (2012). Waldron, Machiavelli, and hate speech. In M. Herz & P. Molnar (Eds.), *The content and context of hate speech: Rethinking regulation and responses* (pp. 345–351). Cambridge: Cambridge University Press.



- Knight, B. (2021). Germany: Hate speech, threats against politicians rise. *Deutsche Welle*, February 9, 2021. Accessed February 10, 2021, from <https://www.dw.com/en/germany-hate-speech-threats-against-politicians-rise/a-56512214>
- Laclau, E., & Mouffe, C. (1985). *Hegemony and socialist strategy: Towards a radical democratic politics*. London: Verso.
- Leigh, I. (2009). Homophobic speech, equality denial, and religious expression. In I. Hare & J. Weinstein (Eds.), *Extreme speech and democracy* (pp. 375–399). Oxford: Oxford University Press.
- McAuley, J. (2020). Disturbing details emerge in beheading of French teacher who showed students Mohammed cartoons. *Stuff*, October 19, 2020. Accessed October 19, 2020, from <https://www.stuff.co.nz/world/europe/300135992/disturbing-details-emerge-in-beheading-of-french-teacher-who-showed-students-mohammed-cartoons>
- Malik, M. (2009). Extreme speech and liberalism. In I. Hare & J. Weinstein (Eds.), *Extreme speech and democracy* (pp. 96–120). Oxford: Oxford University Press.
- Mallet, V., & Murphy, H. (2020). Social media groups under fire in France over Islamist killing. *Financial Times*, October 20, 2020. Accessed October 20, 2020, from <https://www.ft.com/content/b2af266c-5a84-4ddd-8d49-ab41f29a7b06>
- Miller-Idriss, C. (2020a). *Hate in the homeland: The new global far right*. Princeton, NJ: Princeton University Press.
- Miller-Idriss, C. (2020b). Keynote speech, digital conference ‘Hate not found’, Institut für Demokratie und Zivilgesellschaft, Jena, December 11, 2020.
- Molnar, P. (2012). Interview with Kenan Malik. In M. Herz & P. Molnar (Eds), *The content and context of hate speech: Rethinking regulation and responses* (pp. 81–91). Cambridge: Cambridge University Press.
- Mouffe, C. (2005a). *The democratic paradox*. London: Verso.
- Mouffe, C. (2005b). *On the political*. London: Routledge.
- Mouffe, C. (2009). Democracy in a multipolar world. *Millennium: Journal of International Studies*, 37(3). 549–561. doi:10.1177/0305829809103232.
- Mouffe, C. (2013). *Agonistics: Thinking the world politically*. London: Verso.
- Nagel, T. (2002). *Concealment and exposure, And other essays*. Oxford: Oxford University Press.
- Nossiter, A. (2020a). After teacher’s decapitation, France unleashes a broad crackdown on ‘the enemy within’. *New York Times*, October 19, 2020. Accessed October 20, 2020, from <https://nyti.ms/37jBrbC>
- Nossiter, A. (2020b). France, waging a crackdown, honors a teacher and depicts a plot in beheading. *New York Times*, October 21, 2020. Accessed October 23, 2020, from <https://nyti.ms/2HqKSuO>
- Onishi, N., Méheut, C., & Foroudi, L. (2020). Attacks in France point to a threat beyond extremist networks. *New York Times*, November 6, 2020. Accessed February 10, 2021, from <https://www.nytimes.com/2020/11/06/world/europe/france-attacks-beheading-terrorism.html>
- Parekh, B. (2012). Is there a case for banning hate speech? In M. Herz & P. Molnar (Eds), *The content and context of hate speech: Rethinking regulation and responses* (pp. 37–56). Cambridge: Cambridge University Press.
- Royal Commission of Inquiry. (2020). *Ko tō tātou kāinga tēnei. [This is our home.] Report: Royal Commission of Inquiry into the terrorist attack on Christchurch masjidain on 15 March 2019*. November 26, 2020. Accessed December 8, 2020, from <https://christchurchattack.royalcommission.nz/>
- Stephens, B. (2020). The encroachment of the unsayable: Our compromised liberalism has left a generation of writers weighing their words in fear. *New York Times*, Opinion, October 19, 2020. Accessed October 20, 2020, from <https://nyti.ms/31j3BQj>

- Stockmann, D. (2020). China's cat-and mouse game blocking web content no model for EU. *EU Observer*, May 20, 2020. Accessed January 14, 2021, from <https://euobserver.com/opinion/148415>
- Strack, C. (2020). Rabbiner Andreas Nachama zu Facebook: 'Jede falsche Nachricht ist eine zu viel'. *Deutsche Welle*, October 13, 2020. Accessed October 19, 2020, from <https://www.dw.com/de/rabbiner-andreas-nachama-zu-facebook-jede-falsche-nachricht-ist-eine-zu-viel/a-55255574>
- Strossen, N. (2018). *Hate: Why we should resist it with free speech, not censorship*. New York: Oxford University Press.
- Sumner, L. (2009). Incitement and the regulation of hate speech in Canada: A philosophical analysis. In I. Hare & J. Weinstein (Eds.), *Extreme speech and democracy* (pp. 204–220). Oxford: Oxford University Press.
- Trotter, C. (2020). Little's hate speech laws will destroy this government. *Daily Blog*, February 21, 2020. Accessed October 19, 2020, from <https://thedailyblog.co.nz/2020/02/21/littles-hate-speech-laws-will-destroy-this-government/>
- UK Government. (2006). Public Order Act 1986, Part III, Acts intended or likely to stir up racial hatred. Accessed November 9, 2020, from <https://www.legislation.gov.uk/ukpga/1986/64/part/III/crossheading/acts-intended-or-likely-to-stir-up-racial-hatred/2006-01-01>
- Waheed, M., & Hauss, B. (2018). The latest attack on free speech in the Israel-Palestine debate. American Civil Liberties Union, June 5, 2018. Accessed October 19, 2020, from <https://www.aclu.org/blog/free-speech/rights-protesters/latest-attack-free-speech-israel-palestine-debate>
- Waldron, J. (2012). *The harm in hate speech*. Cambridge, MA: Harvard University Press.
- Young, I. (1990). *Justice and the politics of difference*. Princeton, NJ: Princeton University Press.
- Young, I. (2000). *Inclusion and democracy*. Oxford: Oxford University Press.