"Cancel culture on steroids"[1] or a Justified Limitation on Freedom of Expression: New Zealand's Hate Speech Proposals

Pip Dwyer

A dissertation submitted in partial fulfilment of the requirements of the degree of Bachelor of Laws (Honours) at the University of Otago – Te Whare Wānanga o Otāgo

October 2021

[1] David Seymour "PETITION: Stop Labour's Hate Speech Laws" (2021) ACT New Zealand <https://freespeech.act.org.nz/>.

Acknowledgements

To my supervisor, Margaret Briggs, thank you for your invaluable guidance, encouragement and feedback over the course of this year.

To my parents, Helen and Andy, for their constant love and support.

To my friends and flatmates for the memories and laughter, and making this the best 5 years in Dunedin. Special thanks to Joel for proof-reading and being my number one supporter.

Table of Contents

Introduction

Hate speech legislation is extremely controversial, with a careful balancing act that must occur between giving effect to freedom of expression, and protecting targeted groups from the harms of hate speech. As there is no universal definition of hate speech, targeting hate speech towards the intended behaviour is extremely difficult across jurisdictions. There is an alarming rise of hate throughout the globe, manifesting itself in violence and intolerance towards minority groups, making this a pressing issue. This dissertation will explore how this balance is addressed in New Zealand, and how it should ideally be calibrated.

New Zealand has been forced to deal with this issue following the terrorist attack on the Christchurch masjidain in 2019. Following this, the Government established a Royal Commission of Inquiry into the attack, which explored a range of issues relating to hate crime and hate speech. The report released by the Commission[2] identified significant issues in New Zealand's existing legislation, to be discussed throughout this dissertation, and ultimately formed the basis for the New Zealand Government issuing proposals to change the laws regarding hate speech.

This proposal document was released on 25 June 2021, with 6 proposals issued for public feedback and consultation. This dissertation will not discuss all proposals in detail, instead focussing on the proposals that directly relate to the balance of legislating hate speech in a way that allows for freedom of expression, and the difficulty in ensuring that debate and discussion are not limited in an unjustified way. Namely, I will discuss the proposed rewording of both the civil and criminal provisions, as well as the proposal to extend the protection of hate speech laws beyond solely 'colour, race, or ethnic or national origins', as this is the only basis for protection under New Zealand's current legislation.

It must also be noted that this dissertation is targeted towards the provisions currently in the Human Rights Act 1993,[3] as it is these provisions that most firmly target hate speech, in regard to incitement of third parties, and are what the Government's proposals are directed towards. While

[2] *Ko tō tātou kāinga tēnei: Royal Commission of Inquiry into the terrorist attack on Christchurch masjidain on 15 March 2019*, (26 November 2020).
[3] These provisions, namely ss 61 and 131, are contained in Appendix 1 for reference.

there are other pieces of legislation relating to hate speech,[4] these are not within the scope of this dissertation.

Chapter one of this dissertation will discuss the underpinnings of hate speech, in particular the harms of hate speech and the underlying rationales of freedom of expression, and how these harms can justify limiting freedom of expression. It will also detail international agreements New Zealand is party to, as well as the domestic framework for NZBORA limitations, in order to set the basis for the legal analysis of the Government's proposals.

Chapter two will explore the Government's proposals relating to the rewording of both the criminal and civil provisions. It will consider the current wording of hate speech provisions in New Zealand, and how they calibrate the balance between the identified harms of hate speech and freedom of expression principles. It will then examine the proposals and how they seek to recalibrate this balance, using comparative analysis with other jurisdictions to answer the question of whether this has been done in a justified way. I will also discuss potential changes that could be made, based on what has been done in other jurisdictions, in order to better address this balance.

Chapter three will discuss the Government's proposal to extend hate speech provisions beyond race to provide protection to other groups that are subject to hateful speech, with the Government's proposal as it stands looking to include all groups protected from discrimination under s 21 of the Human Rights Act. This chapter will also explore the possibility of including exceptions, using legislation from other jurisdictions as guidance, as a way to ensure that legitimate debate and discussion is not stifled.

Ultimately, these considerations come down to where the line *should* ideally be drawn. This issue faces immense disagreement based upon personal values, with the Government facing significant political scrutiny in their recalibration of the legislation. This dissertation aims to explore these issues in an objective way, considering the ideal way in which the harms of hate speech can be minimised, while still ensuring freedom of expression. The conclusions in each chapter seek to answer this question in relation to the Government's proposals, and while parts of the

[4] These include the Harmful Digital Communications Act 2015, the Films, Videos and Publications Classification Act 1993, the Broadcasting Act 1989 and the Summary Offences Act 1981.

Government's proposals address this issue well, I have made several recommendations as to how the legislation could be calibrated in a more balanced way.

Chapter One: The Harms of Hate Speech and the Underpinnings of Freedom of Expression

Part A: The Harms of Hate Speech

In a 2019 UN strategy document, the key issue of hate speech is summarised: "*There is no international legal definition of hate speech, and the characterization of what is 'hateful' is controversial and disputed*".[5] The starting point for this dissertation must therefore be a clarification of what 'hate speech' is commonly taken to mean, in the legal sense as well as how the term is perceived by society. A distinction can be made between 'soft' hate speech,[6] being speech that is perhaps morally objectionable, but not legally sanctionable, and 'hard' hate speech that meets the legislative threshold. Distinguishing between the two and where the line *should* be drawn is the main focus of this dissertation. I will begin with looking at accepted definitions of hate speech, and a discussion of the harms that arise out of it.

Defining hate speech

The Royal Commission of Inquiry following the Christchurch terrorist attack noted the lack of clarity surrounding hate speech, in particular the ability to reflect the conduct in statutory language. They adopted a definition of "speech that expresses hostility towards, or contempt for, people who share a characteristic"[7] for the inquiry, which summarises the conduct that hate speech is concerned with at a broad level. The 2019 UN Strategy document discussed above describes hate speech as communication that that "attacks or uses pejorative or discriminatory language"[8] towards a group based on a shared characteristic. This definition encompasses conduct such as offensive jokes or derogatory discussion aimed at marginalised groups, which while not necessarily legally sanctionable, alienate groups within our communities and possess harms that are morally reprehensible.

Harms of hate speech

Even without deciding on an exact point where speech becomes legally sanctionable, 'hate speech' in a general sense can have extremely detrimental impacts on both individuals and societies as a

[5] United Nations *UN Strategy and Plan of Action on Hate Speech* (May 2019) at 2.
[6] Stavros Assimakopoulos, Fabienne H. Baider and Sharon Millar *Online Hate Speech in the European Union A Discourse-Analytic Perspective* (1st ed, Springer Nature, 2017) at 4.
[7] Royal Commission Report, above n 2, at 700.
[8] United Nations, above n 5, at 2.

whole. Making groups feel unwelcome and alienated in their communities is a clear harm, and detracts from social cohesion.[9] Despite the old adage of 'sticks and stones', courts have described it as indisputable that emotional harm, leading to social and psychological consequences, can be sustained by the use of words.[10]

In the view of the Supreme Court of Canada, this individual harm, exacerbated by the link between a person's dignity and the respect given to the group within society that they feel they belong to,[11] results in a detrimental impact on a person's self-acceptance. Despite hate speech not being a physical act, research has evidenced that it impacts both an affected individual's mental and physical health.[12] Another commonly cited harm that is attributable to hate speech is the ability it has to silence people and result in disengagement from society, causing individuals to refrain from their normal movement or association.[13]

In addition, there is a second consequence of hate speech described in literature relating to the impact on society as a whole. Expressions of hatred directed at groups within society can conceivably have the outcome of significant social discord, through influencing the views those who hate speech is disseminated towards,[14] even if this influence is not consciously accepted. There is also argument that hate speech harms the public good of an inclusive society, and the value of this should be protected in the interests of society as a whole. Jeremy Waldron advocates for the idea that a justification for the regulation of hate speech is protection of the dignity of individuals.[15] It is this notion of harm to a person's dignity, as opposed to merely offence caused, that Waldron sees as the point at which speech becomes of a legislative concern.

While this can be a difficult line to draw, Waldron states that it should be aimed at ensuring that all members of society are entitled to ordinary social interaction despite belonging to a minority group, as this is what hate speech attacks.[16] The public aspect of this, and the idea stated above

[9] Royal Commission Report, above n 2, at 701.
[10] *R v Keegstra* [1990] 3 SCR 697 at [60].
[11] At [61].
[12] Katharine Gelber and Luke McNamara "Evidencing the harms of hate speech" (2016) 22 Social Identities Journal for the Study of Race, Nation and Culture at 324-341.
[13] Mari J. Matsuda "Public response to racist speech: Considering the victim's story" (1989) 87 Michigan Law Review 2320 at 2327.
[14] *R v Keegstra*, above n 10, at [62].
[15] Jeremy Waldron *The Harm in Hate Speech* (Harvard University Press, Cambridge, 2012).
[16] At 105.

that one of the harms of hate speech is causing individuals affected to disengage from society, appears to contradict the 'marketplace of ideas' theory, one of the justifications for freedom of expression to discussed in Part B of this chapter.

A recent research project in the UK also appeared to find a positive connection between hate speech and violent, physical crime.[17] The investigation focussed on online hate speech,[18] finding a positive association with that and racially or religious motivated aggravated offences. This study advocates for the notion that hate speech should not be perceived as a discrete act, and instead as a process of victimisation. Thus, when considering how the law should deal with hate speech, the focus cannot solely be on the direct harm of hate speech, but also the more tangible crimes that can escalate from it.

Further rationale for hate speech legislation

With the regulation of hate speech being opposed by such strong freedom of expression arguments, the competing rights involved must also be explored, in particular the right to be free from discrimination[19] and the right to equality, both of which are affirmed in various treaties which New Zealand is party to and are at the heart of international human rights law.[20] Freedom of expression should not be used in order to violate or undermine these other individual rights, in particular through the expression of deep-rooted hatred.[21]

Based upon the right to equality, there is an argument that prohibitions on hate speech are not just desirable social policy, but instead a requirement in order for Governments to fulfil their obligations of ensuring that all individuals are able to freely participate in society.[22] Equality is an important democratic value, and therefore the ability for hate speech to undermine this strongly supports government intervention.[23] The Human Rights Commission support this, stating that

[17] Matthew L Williams and others "Hate in the Machine: Anti-Black and Anti-Muslim Social Media Posts as Predictors of Offline Racially and Religiously Aggravated Crime" (2020) 60 British Journal of Criminology 93–117.
[18] In particular, the study focused on tweets in the London area with content expressing hate towards race or religious groups.
[19] New Zealand Bill of Rights Act 1990, s 19.
[20] Frank La Rue *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression* UN Doc A/67/357 (7 September 2012) at [34].
[21] At [37].
[22] Grant Huscroft and Paul Rishworth *Rights and Freedoms: The New Zealand Bill of Rights Act 1990 and the Human Rights Act 1993* (Wellington, Brookers, 1995) at 194.
[23] Katharine Gelber and Adrienne Stone *Hate speech and freedom of speech in Australia* (Federation Press, Annadale (NSW), 2007) at 67.

"regulation that limits speech about race is also symbolic, sending positive messages of inclusion and concern to ethnic minorities and demonstrating a legislative commitment to eradicating racism".[24]

Part B: Freedom of Expression

The key basis for opposing hate speech regulation is centred around freedom of expression, being a fundamental human right in the Universal Declaration of Human Rights and explicitly identified in domestic law through s 14 of the New Zealand Bill of Rights Act. It has become one of the world's most widely accepted rights,[25] recognised for its importance not only on its own but also as a means to enjoy other individual rights.[26] It carries with it numerous justifications against the regulation of hate speech, or at least for regulation to a lower degree, which will now be analysed along with counter arguments, setting a background for the difficulty in legislating hate speech effectively.

A democratic ideal

One of the key arguments made in opposition to hate speech legislation is based upon the idea that freedom of expression is essential for democracy, in order to promote and encourage citizen participation. In the civil libertarian view, it is exactly the expressions that are not approved by the majority that require protection,[27] in order to avoid censorship of criticism towards the Government at the time. A strong, and particularly provoking, discussion involves considering what our society would look like if we did not give effect to freedom of expression, giving rise to a 'slippery slope' argument. This is particularly problematic when considering the implications of a society that does not value freedom of expression, where there are still governments that stifle criticism and dissent regarding challenging social issues.[28]

---

[24] New Zealand Human Rights Commission *Human Rights in New Zealand 2010 – Nga Tika Tangata O Aotearoa 2010* (10 December 2010) at 132.
[25] Dominic McGoldrick and Therese O'Donnell, T "Hate-speech laws: Consistency with national and international human rights law" (1998) 18(4) Legal Studies 453-485.
[26] La Rue, above n 20, at [36]. For example, it is essential in fully realizing the right to freedom of thought, conscience and religion, as in s 13 of NZBORA, and for manifestation of religion and belief, as in s 15 of NZBORA.
[27] Huscroft and Rishworth, above n 22, at 192.
[28] United Nations General Assembly *Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred* UNDoc A/HRC/22/17/Add.4 (11 January 2013) at [5].

However, it is arguable that this justification has little relevance to the context of hate speech. Racist expression is of 'low value' in relation to public discussion[29] and thus critics of the strong civil libertarian approach argue that hate speech regulation does not harm this underlying value of freedom of expression. It is difficult to describe targeted expressions of hatred towards a minority group as a democratic ideal requiring protection.[30] Thus, in the context of hate speech and the individual and group harms involved, there is justification for freedom of expression being limited.

Furthermore, the 'slippery slope' argument can be criticised for setting up a misleading 'either/or' proposition that governments are perpetually hostile and are likely to revert to autocratic tendencies if any qualification is made against freedom of expression.[31] However, it remains important to carefully strike this balance so as to avoid this so-called slippery slope, and thus to ensure legislation only captures expression intended to be targeted, which must be carefully defined and limited.

Marketplace of Ideas

Secondly, the 'marketplace of ideas' justification, attributable to philosopher John Stuart Mill,[32] is commonly used as a means to reject the regulation of hate speech, representing the notion that that all viewpoints should be brought towards a common 'marketplace' in order to advance knowledge and seek truth, with value being placed on differing opinions and criticism. Although to a degree this could be disputed following the same argument as above, that hateful speech communicated at a racial group does not provide substantive value to discussion, the truth-seeking rationale that underpins this argument provides further justification.

Although the marketplace of ideas must be regulated in some way, the argument is that it should not be achieved through the coercive power of government.[33] The idea is that truth can be sought, and separated from untrue or evil statements, through discussion and education in mainstream society. This is deemed to be a better avenue for truth seeking than legal intervention, as governments are more likely to be a product of group biases, and evidence that the most

[29] Huscroft and Rishworth, above n 22, at 194.
[30] Kathleen Mahoney "Hate Vilification Legislation and Freedom of Expression: Where Is the Balance?" [1994] 21 AUJlHRights.
[31] Kathleen Mahoney, above n 30.
[32] John Stuart Mill *On Liberty* (London: Longman, Roberts and Green, 1869).
[33] Eugene Volokh "In Defense of the Marketplace of Ideas / Search For Truth as a Theory of Free Speech Protection" (2011) 97(3) Virginia Law Review 595-603 at 598.

trustworthy consensus is reached when opposing or alternative views are heard, rather than being banned.[34]

However, as discussed in Part A, this can be disputed on the basis that one of the key societal harms of hate speech is its impact on the dignity of the victims and their assurance of an inclusive society. If the marketplace of ideas argument is followed, then there is value in all viewpoints being heard, and if hate speech against minority groups causes the victims to disengage from society, then this justification against the regulation of hate speech is diminished. Further, the idea that the truth will always prevail in the marketplace of ideas, without legal intervention, is potentially dangerous, and historical evidence of racial and religious violence stands as an example of the flaws in this reasoning.[35]

The extent to which this supports government intervention appears to be dependent on the trust that can be placed on them, as opposed to a regime of complete free speech, to seek 'truth'.[36] Letting truth be regulated by society can be dangerous, due to evidence that individuals are disposed to believing dominant views or views that support their needs,[37] which contradicts the proposition that harmful or untrue expressions can be identified and filtered out by society. However, the marketplace of ideas justification recognises that 'truth' is constantly evolving, and thus differing and controversial opinions and expression need to be consistently protected to allow for the development of society's viewpoints.[38]

Personal Autonomy

A third commonly cited justification for the protection of freedom of expression is personal autonomy, being the idea that individuals are entitled to hold and express differing views.[39] Based upon the state's role of deriving their authority from the views of those they represent, they should remain neutral. The personal autonomy justification is rooted in the importance of self-fulfilment

[34] At 597.
[35] Kathleen Mahoney, above n 30.
[36] Kent Greenawalt *Fighting Words: Individuals, Communities, and Liberties of Speech* (Princeton University Press, Princeton (NJ), 1995) at 4.
[37] At 5.
[38] Paul Rishworth, et al. *The New Zealand Bill of Rights* (Oxford: Oxford University Press, 2003) at 309.
[39] Jeffrey Howard, "Should we ban dangerous speech?" (2018) 32 British Academy Review 19 at 20.

and personal development, allowing autonomy in weighing up different views and deciding individual perspective.[40]

However, the subjugation that occurs when the views of those expressing hate speech are prioritised over those targeted must be considered. The personal autonomy rationale faces the same counter-argument as previously discussed, in that if a climate of discrimination is created, targets feel alienated from society, resulting in decreased self-worth and ability to operate with their own autonomy.[41] These harms do not occur with merely offensive speech, therefore provided hate speech laws are directed to the degree that the self-fulfilment and development of targeted groups are impacted, they can be theoretically justified. Abstract theories of personal autonomy, and the marketplace of ideas, do not reflect lived experiences of targeted groups.[42]

Additional jurisprudence supporting hate speech regulation

Critique of the civil libertarian approach, and thus support for the regulation of hate speech, can also be found in legal scholar Horwitz's discussion of rights discourse.[43] One of his key responses to liberal rights discourse was the notion that conceptions of rights were conceived individualistically, and thus operate this way in society, failing to account for and adequately address group harms. In relation to discrimination for example, there tends to be a focus on establishing this in relation to individual sense, and the violation of an individual right, rather than recognising the underlying systematic, structural issue.[44]

This argument applies in relation to hate speech. Although NZBORA 'protects' from discrimination,[45] while simultaneously protecting freedom of expression, criticism can be made that these protections possess an asymmetry. The right to freedom of expression in this context is more adequately protected through individualistic rights, whereas the group harms of hate speech against marginalised groups are not adequately addressed through personal rights as they do not

[40] New Zealand Human Rights Commission *Kōrero Whakamauāhara: Hate Speech An overview of the current legal framework* (December 2019) at 7.
[41] Elizabeth MacPherson, "Regulating Hate Speech in New Zealand" (Wellington 2003) Victoria University of Wellington Faculty of Law at 25.
[42] Mahoney, above n 30.
[43] Morton J. Horwitz "Rights" (1998) 23 HARV. C.R.-C.L. L. REV. 393-406.
[44] At 400.
[45] New Zealand Bill of Rights Act 1990, s 19.

have such an individualistic focus. Therefore, from this jurisprudential lens, there is support for addressing hate speech through legislation.

With particular regard to the different purposes of civil and criminal provisions, this is particularly relevant to discussion of the civil provision, contained in Chapter Two. With preventing discrimination being the purpose of the civil provision, it must be recognised that this is a systemic issue, rather than an individual one, thus influencing the way hate speech legislation should be calibrated in a human rights context.

Part C: International Agreements

Hate speech legislation has been addressed in two key international agreements to which New Zealand is party, thus providing support for limiting freedom of expression in this context.

Article 4(a) of the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD), which formed the origins of sections 61 and 131 of the Human Rights Act 1993, requires states to:

> *"[D]eclare an offence punishable by law all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin*".

Similarly, Article 20(2) of the International Covenant on Civil and Political Rights, states that:

> *"Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law".*

This provides further support for hate speech legislation, and is relevant to the discussion of the proposal to extend the groups protected, as discussed in Chapter Three, as religious based hate speech is not currently incorporated into our domestic law.

Incitement

There is a clear focus here, and thus reflected in our current legislation, on incitement. This relates to the effect the speech has on other people rather than the speech itself, in order to better balance freedom of expression principles and to meet the standard of being a justified limitation. This

threshold negates the misconception that legislation is targeted towards offensive words or 'hurt feelings' as it denotes behaviour to a higher level that influences others.

Those who oppose hate speech laws commonly argue that hate speech legislation should be limited solely to incitement to violence. The National Party has reflected this argument in their response this year to the Government's proposals, stating they will "reverse any attempts Jacinda Ardern's Government makes to criminalise speech beyond the threshold of 'inciting violence'".[46]

This traditional approach was followed by civil libertarians following *Abrams v United States,*[47] promoting that speech should only be regulated when it constitutes action through a 'clear and present danger' test. The traditional approach supports only speech clearly linked with action, such as the hypothetical example given by Justice Holmes[48] as yelling 'fire' in a crowded cinema,[49] as being to the level of requiring intervention. This word/action distinction argues for the notion that speech alone cannot cause harm, although, as discussed in Part A, this can clearly be negated and the harms of hate speech are now recognised to a greater degree.

Furthermore, Eddie Clark notes that this is not how speech is regulated in other areas of the law. He uses examples of both civil defamation, and criminal offences such as blackmail and deception, that are justified based on harms such as a person's ability to interact as part of society and "dignity, autonomy and collective public good".[50] As discussed, it is these kinds of harms that legal theorists argue that hate speech legislation is designed to protect, thus a link to violence should not be a precondition.

---

[46] New Zealand National Party (@NZNationalParty) "National will reverse any attempts Jacinda Ardern's Government makes to criminalise speech beyond the threshold of 'inciting violence.'" <https://twitter.com/nznationalparty/status/1410112520514277377>.
[47] *Abrams v United States* 250 U.S. 6'6 R 620 (919).
[48] *Schenck v. United States 249* US 49 (1918), at 52.
[49] Yelling 'fire' in a crowded cinema was used an example of what would meet this 'clear and present danger' test due to the panic that would be caused, where the words were deemed to be so clearly linked with action so as to justify limiting freedom of speech.
[50] Eddie Clark "Hate speech law must be calibrated carefully" (15 July 2021) Star News <*https://www.odt.co.nz/star-news/star-opinion/hate-speech-law-must-be-calibrated-carefully*>.

Part D: Demonstrable Justification

The overarching legal question is whether limitations on freedom of expression can be 'demonstrably justified in a free and democratic society'[51] by hate speech legislation. All rights in the New Zealand Bill of Rights Act 1990 are protected, but are not absolute. I will briefly discuss the domestic framework for this in order to set the basis for my discussion of the proposals in chapters three and four.

NZBORA Framework

Limitations on rights under NZBORA must be analysed with reference to ss 4, 5 and 6 of NZBORA. In order to determine whether legislation is "demonstrably justifiable in a free and democratic society", as per s 5 of NZBORA, the legislation must limit NZBORA for an objective "sufficiently important" to justify doing so.[52] New Zealand courts have approved the Canadian approach to this question, which requires considering whether there is a "pressing and substantial need for the legislation".[53]

While New Zealand's current civil hate speech legislation has been deemed to fulfil this, with the recognised objective of s 61 being to promote racial harmony,[54] this question will have to be considered in relation to the Government's proposals. In particular, I will discuss in Chapter Three whether there is a 'sufficiently important' objective, and thus a demonstrable need, to justify extension of the of hate speech provisions to more groups, beyond solely race.

Three further questions must be considered in regard to s 5 of NZBORA.[55] Firstly, the limiting measure must be rationally connected to the objective. Following this, it must be questioned whether the limitation on NZBORA impairs the right, in this case freedom of expression, no more than is reasonably necessary to achieve this purpose. While the high threshold set by courts in

[51] New Zealand Bill of Rights Act 1990, s 5.
[52] *Hansen v R* [2007] NZSC 7. The framework for determining whether legislation is 'demonstrably justifiable' as per s 5 of NZBORA is set out in [104]. It is the most recent and authoritative decision relating to NZBORA limitations.
[53] *R v Oakes* [1986] 1 SCR 103.
[54] This was discussed in *Wall v Fairfax New Zealand Ltd* [2018] NZHC 104 at [31]. This purpose was supported with reference to parliamentary commentary relating to its preceding legislation, s 9A of the Race Relations Act 1971, and its introduction in order to meet international commitments under the ICERD.
[55] Again, this is set out in *Hansen v R,* above n 52, at [104].

relation to the current hate speech provisions[56] fulfils this, when it comes to analysing the proposals and the way they seek to recalibrate the balance with freedom of expression, this again will have to be questioned. I will discuss this further in chapter two in relation to the proposed change in wording of the civil provision. The final component of s 5 analysis is an analysis of whether the limitation is "in due proportion to the importance of the objective".

In my discussion of the proposals in chapters two and three, I will reference this framework in my analysis of how these proposals seek to balance freedom of expression New Zealand's hate speech legislation, and ultimately whether the proposals are justified. It must be noted of course, that while it a useful tool in evaluating the legal justification of the proposed legislation, and for the court's in interpreting it, it does not affect its validity as s 4 of NZBORA[57] will prevail.

Application in case law

In terms of how freedom of expression interacts with other NZBORA rights, it must also be noted that the relevant rights in hate speech cases do not always result in a 'classic' conflict of rights.[58] Although s 19 NZBORA, the right to be free from discrimination, is related to hate speech in the sense that the purpose of the provisions is to prevent racial discrimination, it is not directly engaged. NZBORA only applies to acts done as a part of a public function or by New Zealand's legislature, executive or judiciary, as per s 3 of NZBORA.

Thus, when a private individual is prosecuted under a hate speech offence, their obligation in relation to discrimination is solely under the Human Rights Act,[59] thus freedom of expression is the only NZBORA right that is directly engaged. Hate speech offences in this context have therefore been described as a balancing exercise between the right to freedom of expression by private entities or individuals, and the *interest* of the government to protect individuals from harmful speech or discrimination, rather than a *right.*[60]

[56] In *Wall v Fairfax New Zealand Ltd,* above n 54, at [42], s 61 of the Human Rights Act was deemed to be targeted towards speech at the "serious end of the spectrum".
[57] Section 4 denotes that other enactments are not affected "by reason only that the provision is inconsistent with any provision of this Bill of Rights".
[58] *Wall v Fairfax New Zealand Ltd,* above n 54, at [32].
[59] At [33].
[60] At [35].

Chapter Two: Rewording the Provisions

On 25 June 2021, the Ministry of Justice released proposals regarding the incitement of hatred and discrimination in Aotearoa, which ultimately received over 20,000 submissions.[61] At the time of writing this dissertation, the Government has not made any further progress in reforming the relevant provisions. As mentioned, although six proposals have been released, I will not cover them all in detail, and will instead focus on the proposals that most strongly relate to the balance between the identified harms of hate speech and freedom of expression. This chapter will consider the proposed rewording of the criminal and civil provisions.

There appears to be a range of motivations for these changes, including the Royal Commission's recommendations and also a recognised concern that the incitement provisions are seldom used.[62] This has faced some criticism for being contradictory in the sense that they are both trying to insist that the pre-existing high bar will remain, while simultaneously saying they want the proposals to result in more effective law,[63] which shows the pressure in this area with strong competing opinions. This chapter will discuss how the Government seek to recalibrate the difficult balancing act in this area, and whether they have struck this balance right with the proposals as they currently stand.

Part A: The Criminal Provision

The proposed wording of the criminal provision includes 4 elements, as stated on the proposal document:[64]

1. *"intentionally incite/stir up, maintain or normalise hatred*
2. *against any group protected from discrimination by section 21 of the Human Rights Act*
3. *through threatening, abusive or insulting communications, including inciting violence*
4. *made by any means."*

[61] Ministry of Justice "Proposals against incitement of hatred and discrimination in Aotearoa New Zealand: Public submissions are now closed" Justice.govt.nz <https://www.justice.govt.nz/justice-sector-policy/key-initiatives/proposals-against-incitement/>.
[62] Ministry of Justice *Interim Impact Summary: Public discussion document – Proposed changes to the incitement provisions in the Human Rights Act 1993* (18 June 2021) at 5.
[63] Stephen Price "Hating On Hate Speech Laws" (29 June 2021) Media Law Journal <https://www.medialawjournal.co.nz/?p=733>.
[64] Ministry of Justice *Proposals against incitement of hatred and discrimination* (25 June 2021) at 30.

'Hatred'

One of the key changes is to sharpen of the current wording in s 131 of the Human Rights Act 1993 that includes 'hostility', 'ill-will', 'contempt' and 'ridicule' into simply 'hatred'.[65] The Royal Commission recommended this on the basis that it would be more straightforward as to imply "extreme dislike or disgust, including an emotional aversion".[66] The intention behind this is to avoid unclear or imprecise judicial interpretation.

The key case law in this area arose in the application of the civil provision, to be discussed further in Part B of this chapter, where *Wall* held that 'hostility' and 'contempt' were to be interpreted to a standard of "relatively egregious examples of expression which inspire enmity, extreme ill-will or are likely to result in the group being despised".[67] This 'relatively egregious' standard is "not a satisfactory test for the imposition of criminal liability",[68] thus both the Royal Commission and the Government deem this proposal to be less restrictive on freedom of expression, through the removal of wording that has the potential to capture expression by artists or comedians.[69] Legal commentators tend to agree with this, with Graeme Edgeler explicitly stating the proposal would be less restrictive on freedom of expression,[70] and Eddie Clark recognising that the change in wording would be narrower than the existing law.[71]

While hatred is still a relatively amorphous term, it is more commonly used in overseas jurisdictions, such as the UK and Canada where cases have found this wording to be a justified limitation on freedom of expression, and has a clearer meaning denoting expression at the extreme end of the spectrum than the use of four different words that were previously used. Notably, the Supreme Court of Canada in *Keegstra,*[72] in relation to their equivalent criminal hate speech provision, determined that while the provision did limit freedom of expression,[73] it did so in a

---

[65] It must also be noted that the Government is proposing to shift the provision to The Crimes Act 1961, although I will not be discussing this proposal in detail.
[66] Royal Commission Report, above n 2, at 708.
[67] *Wall v Fairfax New Zealand Ltd,* above n 54, at [56].
[68] Royal Commission Report, above n 2, at 708.
[69] Ministry of Justice, above n 62, at 11.
[70] Graeme Edgeler "The Government's Proposed Decriminalisation of Racist Hate Speech" (28 June 2021) Legal Beagle <https://publicaddress.net/legalbeagle/the-governments-proposed-decriminalisation/>.
[71] Eddie Clark "NZ's hate speech proposals need more detail and wider debate before they become law" (27 April 2021) The Conversation <https://theconversation.com/nzs-hate-speech-proposals-need-more-detail-and-wider-debate-before-they-become-law-159320>.
[72] *R v Keegstra*, above n 10.
[73] As protected by s 20(b) of The Canadian Charter of Rights and Freedoms, pt 1 of the Constitution Act 1982, being sch B to the Canada Act 1982 (UK).

justified way through minimal impairment and did so in order to prevent the spreading of hate propaganda, which was deemed to be an important purpose.[74] This is relevant domestically under New Zealand's framework, as discussed Part D of Chapter One, where NZBORA rights should be limited no more than reasonably necessary.[75]

It was clearly stated by the Court that in determining the provision to be a justified limitation, that the Canadian provision "does not suffer from overbreadth or vagueness" as a result of the safeguards present through definitional limits, including the use of 'hatred' which is to be construed to only include the "most severe and deeply felt form of opprobrium".[76] This stands as justification for the sharpening of the wording in the criminal provision to simply hatred, in order to ensure a sufficiently high standard for criminal culpability in order to justify limiting freedom of expression. When looking at the underlying rationale of freedom of expression, as discussed in Chapter 1, expressions to this level of hatred do little to contribute to "the quest for truth, the promotion of individual self-development or the protection and fostering of a vibrant democracy where the participation of all individuals is accepted and encouraged".[77]

Assistance in application of the word 'hatred' can also be found in Canadian decisions, where the Supreme Court in 2013 directed that it should be confined to extreme manifestations "where likely to expose a person or persons to detestation and vilification"[78] This proposal would also ensure domestic legislation aligns with the international covenants previously mentioned in Chapter One, namely the ICERD and ICCPR, which both use the word 'hatred'.

Removal of 'likely to' criteria

The primary concern that appears to have been raised in relation to the criminal provision is the removal of the 'likely to' criteria. The existing criminal provision includes both the requirement for intent and the words or written matter being "*likely to* excite hostility…", whereas the proposed criminal provision would require solely intent. This aspect of the proposal appears to be the result of the Royal Commission saying that they saw the 'likely to' element "as having little or no bearing

[74] *R v Keegstra,* above n 10, at 795.
[75] *Hansen v R,* above n 52, at [104].
[76] *R v Keegstra,,* above n 10, At 700. The facts of this case are an example of what would meet this high threshold, involving a high school teacher communicating extreme anti-Semitic views to his students, for example describing Jewish people as evil and stating they had "created the Holocaust to gain sympathy".
[77] At 701.
[78] *Sakatchewan (Human Rights Commission) v Whatcott* [2013] 1 SCR 467 at 469.

on whether the conduct is sufficiently culpable to justify in a charge"[79] and is therefore unnecessary provided the other criteria are fulfilled.

Stephen Price suggests that this may make the criminal provision too wide, and would be punishing someone for their own thoughts without the need to establish social harm, which may therefore jeopardise its justification on limiting free speech.[80] The harms of hate speech, as discussed in Chapter One, involve both harm to targeted groups as well as society as a whole, which potentially may not exist with the removal of the 'likely to' criteria. Price suggests a likelihood of harm element would provide a suitable middle ground between solely requiring intent and having to prove tangible, real-world harm, such as the eruption of violence, which would be too high of a standard. However, if the proposals were to be implemented as they are, the fact that the Royal Commission has specifically recommended the removal of this element, it may be difficult for the court to interpret the provision this way.

When comparing this proposal to other jurisdictions, in both the UK and Canada there is potential to be criminally convicted solely by establishing intent rather than in combination with a 'likely to" element. In the UK, racial hatred must either be intended to be stirred up, *or* must be likely to be stirred up,[81] and for hatred based upon religion or sexual orientation the criteria is solely intent to stir up hatred,[82] done so in order to set a higher standard to be reached for these characteristics. Similarly, one of the Canadian criminal provisions is based upon *wilfil* promotion of hatred.[83] So, while there is definitely contention on this point, other jurisdictions have indicated intent, without a 'likely to' element may be sufficient to limit freedom of expression.

In fact, the Supreme Court of Canada in *Keegstra* said that the mental element of 'wilful' was sufficiently narrow in scope, being a stringent standard of mens rea[84] which narrows the application of the provision in order to reduce impairment on freedom of expression. In response to the argument that to justifiably limit freedom of expression there must be proof of actual hatred in others, the Court made clear that this would not give proper effect to the psychological harm

[79] Royal Commission Report, above n 2, at 46.
[80] Stephen Price "Hate Speech: a question" (5 July 2021) Media Law Journal <https://www.medialawjournal.co.nz/?p=744>.
[81] Public Order Act 1986 (UK), s 18.
[82] Public Order Act 1986 (UK), s 29B.
[83] Canadian Criminal Code RSC 1985 C 46, s 319(2).
[84] *R v Keegstra,* above n 10, at 700.

suffered by those victim to hate speech, and also the difficulty in establishing a causative link between speech and hatred of a group.[85]

Furthermore, whether speech is 'likely to' cause hatred can be extremely difficult for courts to determine. In *Wall,* the 'likely to' element was discussed in depth, and although related to the civil provision in that case, showed the difficulty in determining *who* exactly must be excited into hostility or contempt.[86] Huscroft summarises the difficulty in this as being that

> *"Determinations as to the likelihood of these harms occurring would seem to depend on the extent to which others are racist, or are considered capable of being influenced by racist expression… Ironically, a determination that the law has been violated ultimately depends on a decision by a human rights body that racist expression was persuasive."*[87]

This discussion shows the difficulty in the 'likely to' standard in this context, perhaps supporting the argument that removing it, and using solely intent, would be preferable in order to justify criminal culpability by avoiding this determination.

Furthermore, another potential justification for the removal of the 'likely to' criteria is that it stands to reason that if a person intends to 'stir up/ incite' hatred, through the use of threatening, abusive or insulting expression, then it is likely to have some sort of effect on the audience. Stephen Price suggests this is likely to be what the Royal Commission was contemplating in their recommendation.[88] He also suggests that the court may interpret the provision in order to read in a likelihood of harm element as implicit in the offence, although this is not entirely clear.

Excluding 'private conversations'

A possible solution may be to follow the Canadian approach, which explicitly includes speech 'other than in private conversations'.[89] It is arguable that by ensuring the intentionally hateful speech is sufficiently public, it would follow on that there is sufficient effect on third parties, and thus the requisite social harm to justify the limitation on freedom of expression. The New Zealand

---

[85] At 776.
[86] While it was stated by the Court that the focus would be on those who are 'susceptible' or 'persuadable' in the analysis of who must be excited into hostility or contempt at [80], they continued to recognise that this cannot be a complete answer due to the varying degrees of 'persuadable' that may exist.
[87] Grant Huscroft "Defamation, Racial Disharmony, and Freedom of Expression" in Grant Huscroft and Paul Rishworth (eds) Rights and Freedoms: The New Zealand Bill of Rights Act 1990 and the Human Rights Act 1993 (Brookers, Wellington, 1995) 171.
[88] Stephen Price, above n 80.
[89] Canadian Criminal Code RSC 1985 C 46, s 319(2).

proposals do not currently possess such a requirement, so perhaps a similar exception should be included to address the concerns raised by Price and other commentators on this point, while simultaneously following the Royal Commission's recommendation and other jurisdictions. This is particularly important for practical purposes in order to avoid people trying to report friends and neighbours for things they have said, which is not the intended purpose of the provision and would address fears of 'Orwellian'[90] laws to at least some extent.

The UK and Wales also recognise the need for a provision of this nature, and currently have a 'dwelling house' exception,[91] where expression contained in a private house, where it is not heard or seen outside of the dwelling, is excluded from being captured by their 'stirring up' offences. With reference to the purpose of criminal law, the UK Law Commission recognises that exceptions of this nature aim to ensure the criminal law is not encroaching on purely private matters, providing further justification for the inclusion of an exception of this nature. However, they also noted that the scope of their exception is "poorly targeted", in the sense that it would exclude conversations to large audiences in a private home, yet would not capture conversations between individuals having a private conservation at another location, such as in an office.[92] Thus, while the UK provides theoretical support for the inclusion of an exception of this nature, I would suggest that the Canadian approach would provide better guidance for New Zealand.

Ultimately, the proposed criminal provision is well balanced, protecting freedom of expression through narrower and clearer statutory language. While there are concerns with the removal of the 'likely to' element, other jurisdictions have indicated that this is justified, although I suggest including a 'private conversations' exclusion may help to ensure that there is requisite harm to justify limiting freedom of expression.

'Maintain'/ normalise

Also included in the proposal to amend the wording of the criminal provision is the inclusion of "maintains or normalises" hatred, in addition to 'incite' or 'stir up'. The consultation document suggests this is to address the concern of 'preaching to the choir' hatred. An example of this would be racist speech at a white supremacist rally that would otherwise meet the criteria of the provision,

[90] Charlie Dreaver "Hate speech laws: ACT, National say proposals a divisive overreach" (25 June 2021) RNZ <https://www.rnz.co.nz/news/political/445548/hate-speech-laws-act-national-say-proposals-a-divisive-overreach>.
[91] Public Order Act 1986, ss 18(2) and (3).
[92] Law Commission (UK) *Hate Crime Laws: A Consultation Paper* (23 September 2020) at [18.252].

but perhaps would not result in successful prosecution as it would not 'incite/stir-up' hatred as those exposed to the speech would already possess those feelings of hatred. The current provisions do not adequately provide for this, where the "verbs 'excite' and 'bring' connote a change in behaviour or thinking",[93] and the same can be said for 'incite' or 'stir-up', thus justifying this addition.

Part B: The Civil Provision

The Government proposals as they currently stand appear to essentially decriminalise hate speech[94] by shifting it to the civil arena. Over a 5-year period between 2016 and 2021, 337 complaints were made to the Human Rights Commission in regard to racial disharmony under s 61 of the Human Rights Act with no successful prosecutions, or even any progression into mediation, which the Ministry of Justice partly attributes to the high threshold set by the provision.[95] One of their stated issues leading to the proposals are that the incitement provisions are seldom used, resulting in the following key changes in the proposals. Again, I will be looking at the proposals relating to the civil provision with the focus on examining how this recalibrates the balance with freedom of expression. I will also discuss the purpose of civil provisions relative to criminal provisions.

The proposals relating to the civil provision are not as developed as the criminal provision at this stage. The Government has agreed in-principle to add the wording "inciting/stirring up, maintaining or normalising hatred" to the existing wording of "excite hostility" and "bring into contempt", but have not gone further than this, for example to sharpen the wording to 'hatred' as in the criminal provision. They are however seeking further consultation on this provision, so there is potential for similar changes to be made at the drafting stage. I will be considering the proposals as they currently stand, as well as considering what changes or additions the Government should consider.

[93] *Wall v Fairfax New Zealand Ltd,* above n 54, at [63].
[94] Graeme Edgeler, above n 70.
[95] Ministry of Justice, above n 62, at 5.

Incitement to discrimination

The proposals also seek to include "incitement to discrimination" in the civil provision. This proposal will be limited to incitement based on discrimination that is already prohibited under Part 2 of the Human Rights Act, for example discrimination in employment, in provision of goods and services and in education.[96] In weighing up the proposals and their potential impact on freedom of expression, I have decided not to discuss this proposal in significant detail, due to the relatively clear scope of the proposal based on existing discrimination provisions, and thus, existing standards for what is considered discrimination. When looking at the purpose of the Human Rights Act, which is to prevent unfair and unequal treatment, this proposal is directed at aligning with this purpose. Furthermore, the inclusion will contribute to fulfilling New Zealand's obligations under Art 4 ICERD and Art 20 (2) ICCPR which both include "incitement to discrimination".

Addition of 'hatred'

Wall indicated that the current use of 'hostility and contempt' in s 61 of the Human Rights Act requires interpretation at a high threshold, essentially to the level of hatred as used in other jurisdictions, applying solely to "relatively egregious examples of expression which inspire enmity, extreme ill-will or are likely to result in the group being despised".[97] *Wall* concerned cartoons in a magazine depicting Māori and Pasifika parents as distracted with 'booze, smokes and pokies' to the detriment of their children's wellbeing, with the Court ultimately deciding that these cartoons did not meet the high threshold denoted by s 61.

However, the Government proposals in their current form look to add hatred on in addition to the current terms. This could suggest that the existing terms, 'hostility' and 'contempt', should be interpreted to denote a lower threshold than they currently do. This is supported by the Ministry of Justice recognising that the proposed civil provision may therefore cover a wider set of behaviours.[98] However, it is not entirely clear what exactly these behaviours are, and as a result "may need other changes to make it clear what behaviour it covers other than incitement of hatred".[99]

[96] Human Rights Act 1993, ss 22, 44 and 57.
[97] *Wall v Fairfax New Zealand Ltd,* above n 54, at [56].
[98] Ministry of Justice, above n 62, at 15.
[99] Ministry of Justice, above n 64, at 21.

Although civil liability does not carry with it the same punishments of a fine or imprisonment, it can result in referral to the Human Rights Review Tribunal and then perhaps the High Court, so there are still civil remedies that can apply,[100] and thus chilling effects on freedom of expression must still be carefully considered. In the view of the Ministry of Justice, by retaining the requirement of incitement of third parties in the proposal, freedom of expression remains adequately accounted for and will allow for robust debate and expression of views.[101] Due to the lack of clarity in the provision in terms of what exactly it is intending to cover, this cannot currently be said with certainty.

By retaining the words 'hostility' and 'contempt', there is potential for the words to be interpreted in a range of ways with a range of meanings, and perhaps in a way that denotes a threshold lower that what can acceptably allow for freedom of expression. Paul Rishworth supports this, saying that "Words such as hostility, ill will and contempt can be interpreted in ways that unduly limit speech that ought to be free".[102]

Canada's approach to 'contempt'

'Contempt' in particular can denote a range of meanings with varying degrees of seriousness. For example, in Canada,[103] it has been defined as "looking down on someone and treating them as inferior".[104] The Court recognised that the word is emotive, and thus can vary between individuals. Therefore, the accepted approach has been to read hatred and contempt together, where hatred colours the meaning and interpretation of contempt,[105] essentially raising the threshold for 'contempt'.

In *Taylor* the majority interpreted 'hatred and contempt' to mean "unusually strong and deep-felt emotions of detestation, calumny and vilification".[106] This approach was emulated in *Wall,* with 'hostility' and 'contempt' being read together to ascertain the 'relatively egregious' standard,

[100] Remedies are listed in s 92I of the Human Rights Act 1993, including various declarations, orders and damages.
[101] Ministry of Justice, above n 62, at 15.
[102] Paul Rishworth *The Right to Freedom of Expression* (A research paper written for the New Zealand Human Rights Commission) in Christine Sanders (ed) *Human Rights Law* (online looseleaf ed, Thomson Reuters) at [HR61.05].
[103] In Canada, the equivalent civil provision was contained in s 13(1) of the Canadian Human Rights Act, RSC 1985, c H-6. It referred to anything "likely to expose a person or persons to hatred or contempt by reason of the fact that that person or those persons are identifiable on the basis of a prohibited ground of discrimination".
[104] *Taylor v Canadian Human Rights Commission* [1990] 3 SCR 892 at 928.
[105] *Nealy v Johnston* (1989) 10 CHRR D/6450.
[106] *Taylor v Canadian Human Rights Commission*, above n 104, at 928.

essentially representing the same level of stringency used in the Canadian approach. What must be questioned however, is whether the proposal to add hatred onto hostility and contempt would consequently lower the severity of speech which would be captured by 'contempt' going forward. The minority in *Taylor* supports there being potential for this, describing 'contempt' as overly broad and stating that it

> "[*d]oes not assist in sending a clear and precise indication to members of society as to what the limits of impugned speech are. In short, by using such vague, emotive terms without definition, the state necessarily incurs the risk of catching within the ambit of the regulated area expression falling short of hatred."*[107]

However, the majority in *Taylor* deemed s 13 of the Canadian Human Rights Act to be justified, even with 'contempt' included in the legislation, as it did not curtail the core values of freedom of expression. This was followed in the case of *Whatcott,*[108] with the Court declaring that this expression prevents the victims from achieving self-fulfilment, the free exchange of ideas and restricts democratic participation.[109] Although the minority held that the proportionality test for what is a justified limitation on freedom of expression[110] could not be met without the requirement of intent, because the legislation is not "rationally connected" to the purpose as it includes speech not intended to discriminate,[111] the majority found that this threshold was met.

Purpose of the civil provision

In reaching this decision, consideration of the purpose of the civil provision, when compared with the criminal provision, was important, and is thus relevant to the discussion of the New Zealand proposals. In response to the criticism that intent should be required to justify limiting freedom of expression, CJ Dickson pointed out that "systemic discrimination is much more widespread in our society than is intentional discrimination", thus requiring intent, rather than focussing on the effects of actions/expression, would defeat one of the key goals of anti-discrimination offences.

---

[107] *Taylor v Canadian Human Rights Commission*, above n 104, at 962.
[108] *Saskatchewan (Human Rights Commission) v Whatcott*, above n 78. This case related to the equivalent Saskatchewan provincial human rights legislation.
[109] At [113]-[114].
[110] Set in *R v Oakes,* above n 53. *Oakes* provides the framework for what is a justified limitation on individual rights under Canadian law.
[111] *Taylor v Canadian Human Rights Commission*, above n 104, at 966-967.

When looking at the purpose of civil human rights legislation, there is support for this, where legislation is not about punishment or stigmatisation of those who discriminate, instead concerned with preventing discriminatory conduct and protecting victims from this.[112] In this sense, the purposes and consequences of civil provisions are different than that of criminal ones, and therefore must undergo different tests in determining whether the legislation is a justified limitation on freedom of expression. With this purpose in mind, perhaps an interpretation of 'contempt', lower than the standard of 'hatred', could be justified in the sense that the same discrimination and associated harms can result.

It must be noted that the Canadian Government decided to repeal s 13 of the Canadian Human Rights Act in 2013. The timeline of this repeal is interesting, as it "highlights how the government advanced its own agenda as it relates to the proper balance to be struck between freedom of expression and hate speech legislation without awaiting the Supreme Court's pronouncement on the same issue".[113] The Bill was presented before the House of Commons prior to the *Whatcott* decision being released, with the conservative majority passing the Bill, effectively promoting unfettered expression and discarding the legal analysis of Canada's highest Court. Despite the Supreme Court stating the importance for society of the civil hate speech provisions, the conservative majority in the Government effectively disregarded this.

This shows in balancing between the harms of the public dissemination of hate speech and freedom of expression, as both the judiciary and the Government were undergoing the same recalibration at different times with different outcomes. However, the decision of the Government at the time appears to be politically motivated, and subsequently, the Justice Committee has recommended the provision be reinstated to address the continued presence of hateful speech.[114] What the Canadian approach to the civil provision indicates is that from a legal standpoint, civil provisions for hate speech can be justified, including with the wording of 'contempt' in the view of the majority in *Taylor,* although political decisions may result in a different outcome.[115]

---

[112] Lauren E Scharfstein "The Hate Speech Debate: The Supreme Court, the Federal Government, and the Need for Civil Hate Speech Provisions" (2019) 19 Asper Review of International Business and Trade Law 375 at 382.
[113] At 393.
[114] House of Commons (Canada) *TAKING ACTION TO END ONLINE HATE: Report of the Standing Committee on Justice and Human Rights* (June 2019).
[115] As such, although there is remaining provincial civil hate speech provisions in many Canadian provinces, there is currently no federal civil provision.

Application in New Zealand

While the Canadian Supreme Court approach in both *Taylor* and *Whatcott* provides support for civil hate speech provision providing a justified limitation on freedom of expression, even with the inclusion of 'contempt' provided it is interpreted to a high threshold, the concern that remains with New Zealand's current proposal is the lack of clarity. It is difficult to know whether the court would continue to read the words in conjunction, and interpret s 61 to a high standard, or whether the court would interpret contempt, as well as hostility, at a lower threshold to reflect the intention behind the legislation. As it stands, it is not clear to the public what would be prohibited by this section. Edgeler raised this concern in response to the proposals, in that the Government has not been clear about what kind of conduct they intend to ban and have struggled to respond to hypotheticals given to them.[116]

Based on the plain meaning of contempt, or perhaps if it were to be interpreted to mean 'looking down on someone' as per the definition suggested in *Taylor,* it could capture expressions far less extreme in nature. Taking the case of *Wall,* it is possible that if a lower threshold for contempt was used, the outcome may have been different. It is not overly difficult to imagine that the cartoons in question could incite third parties to look down on Māori and Pasifika people and treat them as inferior based on the cartoons.

'Contempt' and similar words used in other contexts may be used to aid the court, for example comparisons can be drawn between contempt in this context and the use of the 'discrimination and denigration standard' in the Broadcasting Act[117] which is cast in similar terms.[118] In *Day v NZME Radio Limited,* the Broadcasting Authority justified the limitation of freedom of expression and upheld the complaint on the basis that of the comments made by a radio presenter having "significant potential to cause harm, through distress and denigration",[119] which requires a high level of condemnation. This case provides potential guidance for the courts if the proposals were to be implemented as they stand, as to how contempt could be interpreted at a lower standard. Harm through distress and denigration is significantly lower than the current interpretation of

[116] Graeme Edgeler, above n 70.
[117] Broadcasting Act 1989, s 21(e)(iv).
[118] Christine Sanders (ed) *Human Rights Law* (online looseleaf ed, Thomson Reuters) at [HR61.05].
[119] *Day v NZME Radio Ltd* [2019] NZBSA 25 at [18].

contempt seen in *Wall,* where in combination with hostility, was treated essentially to the same level as hatred.

Although it is not clear exactly how the court would interpret and apply s 61 if it were implemented to reflect the proposals, it is ultimately this lack of clarity that is concerning. The Court in *Wall* made clear that there needs to be "broad space" to "express views which may offend, shock or disturb required in a free and democratic society".[120] The Tribunal made clear that the text and purpose of the statute denoted a high threshold, and this was agreed upon by the Court. If these proposals are with the purpose of lowering the threshold, this could perhaps change. Again, the key concern is whether this will leave room for discussion, debate, criticism and even satire which are important in a free and democratic society, and perhaps do not have the extent of harm as inciting pure hatred.

Rationale for lowering the threshold

A secondary question therefore arises, in that it must be considered whether the threshold *should* be lowered. In determining how the balance with freedom of expression should ideally be calibrated, the harms of hate speech and the rationale underpinning freedom of expression, as discussed in Chapter One, must again be referred to. As mentioned, the Ministry of Justice stated a concern regarding the high level of complaints regarding racial disharmony, none of which progressed further to mediation. This lack of prosecutions was raised as a concern by the United Nations Committee on the Elimination of Racial Discrimination in their 2017 concluding observations.

Furthermore, extending the speech covered by the provision is suggested by the Ministry of Justice as being appropriate and justified in order to address rising extremism in New Zealand society, perhaps by capturing online speech associated with the Christchurch terror attack in March 2019.[121] The Ministry of Justice recognises the difficulty in assessing exactly how often inciting speech actually occurs, although evidence from focussed engagements with individuals from groups who 'experience hate speech' formed the basis of the notion that there are issues present in the current provisions.

[120] *Wall v Fairfax New Zealand Ltd,* above n 54, at [56].
[121] Ministry of Justice, above n 62, at 5.

Consequently, perhaps there is valid reasoning for lowering this standard. The key however is in ensuring freedom of expression is adequately allowed for and this requires greater clarity on what the proposals are trying to protect. As per New Zealand's NZBORA framework,[122] discussed in Chapter One, limitations must not impair freedom of expression more than what is reasonably necessary to achieve the purpose of the legislation. If the purpose of the legislation is to reduce social disharmony and resulting discrimination, the fact there are so few productive outcomes resulting from the current legislation may suggest further protections are needed, and are thus reasonable in the circumstances. In saying this however, the key issue remaining is the lack of clarity present in the proposals as they stand, as to ensure that freedom of expression is not impaired more than reasonably necessary the drafting must be sufficiently narrow and clear in scope.

Further clarity is essential in order to avoid negative implications, and practical considerations need to be contemplated, such as the potential increased workload of the Human Rights Commission in dealing with complaints where it is unclear what exactly the conduct is that is intended to be covered. Furthermore, in order to avoid chilling effects on freedom of expression, clarity is essential to avoid self-censorship, where expression may be stifled due to a fear of repercussions, where it would not be speech justified in limiting. With one of the key underpinnings of freedom of expression being personal autonomy, this is concerning.

Application by the Court

A lack of clarity in human rights provisions should be avoided to the greatest degree possible, with the Supreme Court observing that "the more vague the purpose and meaning of an enactment, the less protection for human rights".[123] This places greater interpretive responsibility on the courts. In interpreting the existing civil provision, the Court in *Wall* determined that 'contempt' and 'hostility' are 'conceptually elastic', thus the Court applied the *Moonen*[124] framework to ascertain the meaning intended by Parliament.[125] This approach dictates that the "point at which a tenable

[122] *Hansen v R,* above n 52, at [104].
[123] *Morse v Police* [2011] NZSC 45 at [16].
[124] *Moonen v Film and Literature Board of* Review [2000] 2 NZLR 9.
[125] While the *Hansen* framework, discussed in Part D of Chapter 1, is the most recent and authoritative decision relating to NZBORA limitations, in some cases, where wording is "conceptually elastic", as in producing continuum of meanings, as opposed to two distinct interpretations, the *Moonen* framework may instead be applied. This was discussed in *Wall v Fairfax New Zealand Ltd,* above n 54, at [40].

meaning ceases to limit or least limits the right or freedom may well represent the appropriate point at which to fix the meaning".[126]

In the application of this in *Wall,* the Court's interpretation resulted in a high threshold, targeted towards racist speech at the "serious end of the spectrum". The Tribunal recognised that "an overbroad interpretation of s 61(1) could well unreasonably limit freedom of expression".[127] In relation to the proposals, it is therefore likely that courts will interpret provisions that have a continuum of meanings to denote a high threshold, in order to limit the impact on freedom of expression. Therefore, despite the addition of hatred to the civil provision and the expressed intention of lowering the threshold, courts may continue to interpret the civil provision to the same high threshold as currently present, due to a lack of clarity in terms of other conduct that would be captured by the provision.

Conclusion

I suggest there are two possible options in order to achieve greater clarity. Firstly, as the proposals are open to consultation, perhaps after the public submissions process there is potential for the civil provision to streamline 'hostility' and 'contempt' to 'hatred', as was done with the criminal provision. This would address the concern discussed regarding the interpretation of contempt and the potential implications of a lack of clarity in this provision, in terms of the need for 'broad space' for debate and criticism. However, this would not extend the protection offered by the provision, which the Government has indicated they want to do.

Alternatively, exceptions or 'freedom of expression protections' could be used as a means to clarify what conduct is intended to be captured by the provision and to ensure legitimate discussion is not included, to be explored in Part B of Chapter Three. As I will discuss however, overseas jurisdictions tend to use these solely in relation to criminal provisions, thus it is perhaps unlikely that these will be used by New Zealand in relation to the civil provision.

If the proposal is implemented without any further changes, with 'hatred' added alongside the existing wording of 'hostility' and 'contempt', Canadian case law indicates this will be legally justified. However, it is possible the courts will continue to read the words together to denote a high threshold, essentially to the level of hatred, and thus also may not extend the protection

[126] *Wall v Fairfax New Zealand Ltd* [2017] NZHRRT 17.
[127] At [192].

offered by the provision. There is a clear difficulty here in achieving both the Government's objective of a more effective provision, while still maintaining a high threshold so as to allow for freedom of expression, perhaps suggesting a trade-off will need to occur in the implementation of this legislation.

Chapter Three: Extending the Protection of the Incitement Provisions to More Groups

Currently, both the criminal and civil incitement provisions apply to groups solely based upon "colour, race, or ethnic or national origins". The Government has recognised that hateful speech is also directed at other groups,[128] and thus are looking to extend the protection of the incitement provisions. They are seeking consultation on other groups that experience hateful speech, with the potential groups being those currently protected from discrimination, listed in s 21 of the Human Rights Act.[129]

This wide extension was agreed in-principle by Cabinet.[130] The benefits of extending protection to all groups listed in s 21 of the Human Rights Act is that it would provide for consistency with the Act's focus on discrimination, as incitement is considered to be a form of discrimination, as well as clearly identifying and providing for the groups protected.[131] However, there is also significant controversy and disagreement in regard to this proposal, an issue that other jurisdictions have also faced when looking to amend their own hate speech legislation. There are arguments both in favour of 'blanket' protection of all groups, as well as counter arguments, which are likely to be a significant issue for the Government to consider in their review of submissions.

Part A: Extension of the Groups Protected

Theoretical arguments

It must be considered what the impact on freedom of expression would be if the incitement provisions were to apply to protect all groups based on the characteristics listed in s 21 of the Human Rights Act. One of the key arguments against this, is that some of these groups have greater value in terms of the democratic importance of being able to express controversial or critical views, for example religion and political opinion. A related argument is that there is no demonstrable need for this legislation to be extended.

As discussed in Chapter One, racist speech has little democratic importance,[132] and thus on a theoretical level does not threaten this freedom of expression rationale, therefore justifying

[128] Ministry of Justice, above n 64, at 17.
[129] The prohibited grounds of discrimination are sex, marital status, religious belief, ethical belief, colour, race, ethnic or national origins, disability, age, political opinion, employment status, family status and sexual orientation.
[130] Ministry of Justice, above n 62, at 10.
[131] Ministry of Justice, above n 62, at 11.
[132] Mahoney, above n 30.

government intervention to a higher degree. However, for other characteristics, the same rationale perhaps does not apply to the same extent in relation to the fundamental values of freedom of expression. The 'democratic ideal' justification for freedom of expression is formed on the basis that discussion of ideas and opinions, even if they are opposite to the view of the majority, are important to ensure all individuals have the right to participate in public discourse and society. In relation to political opinion and religion for example, there is broader scope for discussion that would contribute to a democratic society, for example debating matters of policy, such as immigration, or religious ideas such as abortion.

However, as I will discuss, there is a strong counter argument that provided the threshold of the provisions is set high enough and with sufficient clarity, extending the protected characteristics may be justified. The broader harms of hate speech are not limited to race, in terms of psychological harm of the targeted groups. Furthermore the other core values of freedom of expression also face the same threat, irrespective of the group targeted, in terms of the way that the speech can result in members of the targeted group refraining from participating in society and exercising their own personal autonomy. The inclusion of exceptions, which I will detail later in this chapter, may help in ensuring that the scope of New Zealand's hate speech do not limit legitimate debate and discussion.

A further point that must be discussed is that in order for a limitation on freedom of expression to be justified, it must be established that there is a sufficiently important purpose for implementing the legislation,[133] meaning there must be a demonstrable need for it. In regard to the proposal to extend the protection of hate speech legislation to more groups, there is a lack of evidence as to the extent of the problem for each characteristic in New Zealand. There are no conclusive statistics as to how many annual occasions of incitement speech towards each characteristic there are, at least in part due to the fact there is no legislation relevant to them. The Ministry of Justice has recognised this issue in our reform process, emphasising that it is "difficult to assess how often inciting speech occurs",[134] thus relying on engagement and focussed consultation with affected groups.

[133] *Hansen v R,* above n 52, at [104].
[134] Ministry of Justice, above n 62, at 6.

Despite this, other jurisdictions have dealt with the same issue, and thus can provide some guidance. In Scotland's recent reform process, while it was recognised that religion was deemed to have the strongest basis for extension of protection, Lord Bracadale stated that, using age and disability as examples, that the *"fact that it is less likely that people may commit offences of stirring up of hatred against people on the grounds of their disability or age did not mean that stirring up of hatred on these grounds should not be criminalised*".[135] If an individual's speech is intended to stir up hatred, and reaches the same threshold, this approach suggests that even though it may be rarer, it should still result in the same legal consequences.

Guidance from the UK

England and Wales can also be looked at for guidance. The UK Law Commission has discussed the matter of extending the protection of their equivalent 'stirring up' offences on two occasions.[136] In 2014, they considered the extension of the provisions to include disability and gender, and concluded that while extension could be justified in principle, it could not be demonstrated that there was a practical need for the extension.[137] At this point in time, the UK Law Commission found that there was not sufficient evidence that inciting speech was occurring on a significant scale.[138] Thus, despite recognising the theoretical gap, where the law does not protect against speech that would otherwise meet the high standard of incitement, thus making the legislation inadequate, the UK Law Commission did not recommend to extend the provisions.

However, in their 2020 report, they reached a different outcome, based upon increased evidence of hate crime offences relating to disability and transgender groups.[139] Despite noting issues with these statistics, notably the fact that hate crime offences do not necessarily correlate with stirring up offences and the improvement of recording practices by Police, they decided it was legitimate to read across this evidence. For this reason, they determined there was now a demonstrable need for the extension of the legislation, based upon increased practical justification.[140]

What this means for New Zealand, is that while the same theoretical justification may apply, it is the practical justification that is likely to be lacking. The Ministry of Justice has recognised this

[135] The Scottish Parliament *Hate Crime and Public Order (Scotland) Bill Policy Memorandum* (April 2020).
[136] They released a report on this in 2014, then again in 2020.
[137] Law Commission (UK), above n 92, at [18.213].
[138] Law Commission (UK) *Hate Crime: Should the Current Offences be Extended?* (May 2014).
[139] Law Commission (UK), above n 92, at [18.216].
[140] Law Commission (UK), above n 92, at [18.223].

concern, stating that there is "no baseline data on the occurrence of inciting speech".[141] It is therefore difficult to conclusively establish that there is a need for the legislation. At this point, the Ministry of Justice has based the proposals largely on focussed engagement with communities deemed to be affected, consultation with organisations and national and international research,[142] which provides some support for extension, but is arguable whether this fully represents a 'sufficiently important purpose' to justify curtailing freedom of expression for each characteristic.

While the UK's approach of reading across evidence from hate crimes may also provide some guidance, conclusive statistical evidence in this area is also lacking in New Zealand. However, a 2018 Netsafe survey determined that perceived online hate speech was more prevalent against minority ethnic groups, people with disabilities, and non-heterosexual individuals, and that religion as the most common reason to be targeted at a rate of 24%.[143] While these statistics do not directly show the amount of speech to the level of legal incitement, evidence of this nature provides support for extending the protected groups to some degree.

What New Zealand must be careful not to do is to implement law that does not have a justifiable basis, as that would be to "use the criminal law as a tool of social research".[144] As the proposals are currently at the consultation stage, it is possible the New Zealand Government may need to conduct further research across all proposed protected groups in order to demonstrate a pressing need. While discussion of all proposed characteristics would not be possible in the length of this dissertation, I will look to the research done by other jurisdictions later in this chapter to discuss possible approaches for New Zealand through comparative analysis, as well as looking at other potential responses raised by legal scholars.

Responses

In response to the issues discussed above, Graeme Edgeler argues that provided the law is drafted tightly enough to focus on the worst type of speech, at the far end of the spectrum, and not focused

[141] Ministry of Justice, above n 62, at 21.
[142] Ministry of Justice, above n 62, at 6.
[143] Netsafe *Online Hate Speech: A survey on personal experiences and exposure among adult New Zealanders* (November 2018) at 16.
[144] Law Commission (UK), above n 138, at [7.145].

on debate or criticism (even if it may be harsh or leading to ridicule) then extending the groups protected should not be an issue.[145]

His response to the argument discussed above, that groups beyond race such as political opinion and religious expression have greater democratic importance, is to emphasise that provided the law is calibrated in a way that clearly only includes speech that inspires hatred in others of people with that characteristic, as opposed to hatred of the characteristic itself, then the proposed law should not be an issue.[146] Although this is a difficult distinction in reality, if the proposals are going to extend the groups protected by hate speech laws to this extent, then there is further importance placed on ensuring that the criminal and civil provisions set a high enough threshold to ensure limits on freedom of expression are justified. The discussion of exceptions in Part B of this chapter can aid in ensuring the clarity of these provisions.

He draws a parallel between the Christchurch terror attacks in 2019 with the terror attack in Norway in 2011, where those targeted were members of the Worker's Youth League, which is affiliated with a political party. Edgeler argues that if part of the intent of hate speech legislation is to diminish the likelihood of extremism in our society, then drawing the line at political opinion may not support the legislation's intended purpose. It seems illogical to protect one group in society from hateful speech, and allow it based on another, provided it is to the same extreme level.[147]

In relation to the above discussion of whether there is demonstrable justification for extending the provisions, by targeting a high threshold there is support for this based on the idea that at this level the harms of hate speech, discussed in Chapter One, apply across all characteristics. The goal of social cohesion and an inclusive society where all groups can participate without fear clearly applies to groups other than those defined by their ethnic background. The same psychological harms can arise, and as shown by Edgeler's discussion of the Norway terror attack, can have the same links to physical crime. While perhaps for some characteristics, the high standard of

[145] Graeme Edgeler (11 July 2021) "Political opinion and the proper scope of hate speech laws; a post in honour of John Campbell" Legal Beagle <https://publicaddress.net/legalbeagle/political-opinion-and-the-proper-scope-of/>
[146] For example this would mean inciting hatred towards Christian people, as opposed to Christianity itself.
[147] Edgeler, above n 145.

incitement of hatred may be less likely to be reached, it is worthwhile having safeguards in place to protect against this.

Scotland's approach

When looking at other jurisdictions and the way they have dealt with hate speech laws, there have been similar issues raised with proposals to extend the protection they offer. New Zealand's current situation is comparable to that in Scotland, where in 2019 it was proposed that 'stirring up' offences be added to protect a broader range of groups than race. Upon consultation there was an extremely mixed response, which is likely to foreshadow what can be expected by the responses to our proposals.[148]

The Scottish reforms stemmed from Lord Bracadale's investigation and recommendations,[149] where although religion was deemed to have the strongest case to justify protection, the interest of legal 'parity' justified extending the protected characteristics for the stirring up offences more broadly.[150] Upon consultation, many respondents echoed Lord Bracadale's reasoning in that doing so would avoid a 'hierarchies' of victims outcome, and recognised the seriousness of the harms of hate speech, identifying a broader range of individuals and communities targeted.[151] This provides support for New Zealand extending protection to all groups listed under s 21 of the Human Rights Act.

On the other hand, consultation did unearth opposition to the extension, with concerns raised about the impact on freedom of expression, with particular regard to the difficulty in distinguishing between 'threatening and abusive' expression as opposed to disagreement or debate. A common theme of the responses indicated a concern that the proposals would result in unjustified complaints being lodged for ulterior motives, and suggested self-censorship would contribute to negative impacts on freedom of expression. These concerns are thus relevant to New Zealand's proposals, with the potential that they will be raised through the current consultation process. There was particular opposition from faith-based groups who believed widening the stirring up offences

[148] The Scottish Government, *One Scotland: Hate has no home here – Consultation on amending Scottish hate crime legislation Analysis of responses: Final Report* (June 2019) at 4.
[149] Lord Bracadale *Independent Review of Hate Crime Legislation in Scotland Final Report* (31 May 2018).
[150] The Scottish Government, above n 148, at 83.
[151] At 84.

would shut down debate in areas such as same sex marriage and religion, and impact the ability of religious teachings to occur.[152]

This fear of chilling public debate and discussion has been discussed in Canada by the Supreme Court in *Whatcott*. The Supreme Court criticised the Court of Appeal decision, who declared that a higher level of tolerance was required for 'public policy debates', and said that labelling the expression as part of this discussion or debate should not end the inquiry of whether the expression is hate speech, referring to historical rhetorics that demonstrate some of the most damaging expression could fall within this category of debate. The Court said it does not 'chill' the public discussion, just ensures that extreme levels of expression, as to the threshold of "detestation and vilification" are not included in this debate.

Ultimately, following the consultation process, the Hate Crime and Public Order (Scotland) Act 2021 came into force in April, with Parliament creating new stirring up offences on the basis of age, disability, religion or, in the case of a social or cultural group, perceived religious affiliation, sexual orientation, transgender identity and variations in sex characteristics.[153] After weighing up the conflicting views discussed above, the Scottish Government discussed the harms of hate speech on the groups targeted, as well as the social atmosphere it can create, to ultimately implement a wide extension of groups protected by their 'stirring up' offences.[154] Lord Bracadale's discussion of legal parity was an important factor in this.

The way in which the Scottish Government have implemented these new offences can provide insight to New Zealand on how this can be done while still giving effect to freedom of expression, with one of the key additions that could be considered to be included in our proposals being a 'freedom of expression' provision or exception. There has not yet been any cases to test the application of these provisions, although similar provisions in England and Wales provide some guidance on this.

[152] At 85.
[153] Hate Crime and Public Order (Scotland) Act 2021, s 4(3).
[154] The Scottish Parliament, above n 135.

Part B: Exceptions

Guidance from the UK

Scotland has followed the approach of England and Wales, who have specific protections for freedom of expression in their stirring up offences related to sexual orientation and religion. Currently, England and Wales solely have stirring up offences in regard to these characteristics, as well as race, but are also looking at an extension of their provisions. These freedom of expression protections have been crafted in order to clarify the scope of their stirring up offences, identifying expression that should not be limited in a free and democratic society. They aim to address concerns that hate speech provisions will stifle legitimate discussion and have a chilling effect on freedom of expression.

The UK Law Commission describe the common thread of these exceptions as:

1. *"clarifying that the law applies to hatred against persons, not against institutions or belief systems;*
2. *clarifying that criticism of behaviour is permitted; and*
3. *maintaining a space for discussion of public policy on potentially controversial issues".*[155]

Scotland's recent reform looked to these UK and Wales provisions for guidance. Both jurisdictions have treated religious expression differently to the other characteristics, likely due to the increased controversy and opposition surrounding faith-based groups. Scotland's freedom of expression provision states that for all characteristics aside from race and religion, "*behaviour or material is not to be taken to be threatening or abusive solely on the basis that it involves or includes discussion or criticism of matters relating to*"[156] the characteristic. For religion, the protection is more specific and includes:

"*(b) discussion or criticism relating to, or expressions of antipathy, dislike, ridicule or insult towards—*

(i) religion, whether religions generally or a particular religion,

(ii) religious beliefs or practices, whether religious beliefs or practices generally or a particular religious belief or practice,

(iii) the position of not holding religious beliefs, whether religious beliefs generally or a particular religious belief,

[155] Law Commission (UK), above n 92, at [18.274].
[156] Hate Crime and Public Order (Scotland) Act 2021, s 9.

(c) proselytising, or

(d) urging of persons to cease practising their religions."[157]

In Scotland's reform process, consultation specifically asked for feedback regarding the inclusion of a provision of this nature and the vast majority of respondents were in support of this, with 91% (955 out of 1047 respondents) being in favour of a freedom of expression protection.[158] Respondents in Scotland commonly referred to the 'balance' with freedom of expression and the difficulty in striking this appropriately,[159] with provisions of this nature helping to calibrate this clearly.

The UK Law Commission, when discussing the provisions in England and Wales, also stated that they saw the provisions as "necessary to provide clarity and reassurance that legitimate discussion is not caught by the offences".[160] In discussing whether the stirring up offences should be extended in England and Wales, they were particularly concerned following *Miller v College of Policing,*[161] where Police treatment of the complainant following their tweets being reported as transphobic was held by the High Court to be a breach of their right to freedom of expression.[162]

It is important to clarify that there is often a misconception about the type of conduct that will be covered by hate speech provisions. For example, 'deadnaming', as in referring to transgender people by their former name, is unlikely to meet the threshold of stirring up hatred,[163] despite being highly offensive to transgender people. By the same reasoning, discussing Laurel Hubbard's ability to compete at the Olympic games is also unlikely to meet this threshold.

The judgment in *Miller* referred to evidence of a tendency of those involved in an intolerance of different viewpoints even when those viewpoints "are not grounded in hatred, bigotry, prejudice or hostility, but are based on legitimately different value judgments, reasoning and analysis, and form part of mainstream academic research".[164] It is this desire to avoid censorship of public debate, and the consequent reporting of legitimate discussion to authorities, that led to the Law

[157] Hate Crime and Public Order (Scotland) Act 2021, s 9.
[158] Scottish Government, above n 148, at 82.
[159] At 93.
[160] Law Commission (UK), above n 92, at [18.270].
[161] *Miller v College of Policing* [2020] EWHC 225.
[162] Under the European Convention on Human Rights.
[163] Law Commission (UK), above n 92, at [18.222].
[164] *Miller v College of Policing,* above n 161, at 250.

Commission accepting that the freedom of expression clauses should be implemented for all new characteristics protected by the incitement provisions.

The New Zealand proposals as they currently stand do not have any provision for freedom of expression, thus this may be something that the Government considers following the consultation process. In respect of the criminal provision, a protection of this nature reflects the seriousness of the consequences for breaching the provision, helping to set and clarify the high threshold.

Also, although the UK provisions are solely in respect to their criminal provisions, a freedom of expression protection of this nature may also help calibrate the balance with freedom of expression more appropriately in the civil arena. As discussed, the proposals as they currently stand intend to add hatred in addition to the existing standard of 'hostility and contempt'. Perhaps by having an explicit freedom of expression protection, it could allow for the lower threshold which the Government is wanting, while still making clear the type of conduct, such as discussion and criticism, are not within the scope of the provision.

Potential issues

The potential issues with a provision of this nature must be discussed however, with the UK provision being criticised for making the incitement provisions devoid of any practical impact, with it being difficult in reality to distinguish between an attack on a belief system and an attack on those who adhere to it.[165] While it was noted by the Commission that 'stirring up' offences in regard to sexual orientation and religion had fewer successful prosecutions than racial hatred, it was difficult to attribute this to the free speech provisions as opposed to other elements raised the liability threshold such as establishing intention.[166]

There is a further issue raised in the UK Law Commission's discussion of England and Wales' current protections and their discussion on whether the groups protected should be extended. They noted that their free speech exceptions have the potential to create difficulty where groups may be classified as based upon both religion and race, and thus can be subject to hatred based on both characteristics. As a result, they note that it could result in a situation where there is an existing

[165] Law Commission (UK), above n 92, at [18.265].
[166] At [18.265].

defence with the freedom of expression protection in relation to religion, but there will be no defence for the stirring up of racial hatred.

This is particularly relevant in a New Zealand context, where the only prosecution under s 25 of the Race Relations Act, the provision that preceded s 131 of The Human Rights Act, was concerned with whether Jewish people were protected by "ethnic origin". Despite the Court of Appeal in *King Ansell v Police*[167] emphasising that the statute was "not directed at discrimination based on religion",[168] it was defined relatively broadly as to mean "a segment of the population distinguished from others by a sufficient combination of shared customs, beliefs, traditions and characteristics"[169] to include Jewish people.

Furthermore, there is also argument that the current provisions, by only protecting racial hatred, create anomalies, as it treats Jewish and Sikh people, with Sikhs also being held to fall within "ethnic origins" in a House of Lords decision,[170] differently from other religions. While this provides support and justification for extending the protection of incitement provisions to include religious hatred, it must be ensured that this anomaly does not transfer into the applicability of freedom of expression provisions. This is a concern with the current UK and Wales legislation, where words and behaviour against Muslim people could potentially be protected under their freedom of expression provision relating to religion, even if it could otherwise be established as intentionally stirring up hatred, whereas similar behaviour towards Jewish people would not be protected as there is no such provision relating to racial hatred.[171]

The conclusion reached by the UK Law Commission in response to these concerns would be to have their freedom of expression protections apply across all groups, in order to avoid ad hoc application. For example, instead of the protection of freedom of expression in relation to religion applying solely to religious hate speech, it would apply to all of the incitement provisions, noting that in reality in most cases the protection would only be relevant to one form of hatred.[172] This is important to note if NZ is to adopt these protections, and a similar approach should be followed.

---

[167] *King-Ansell v Police* [1979] 2 NZLR 531.
[168] At 541.
[169] At 543.
[170] *Mandla v Dowell Lee* [1983] 2 AC 548 (HL).
[171] Law Commission (UK), above n 92, at [18.271].
[172] At [18.275].

Other Exceptions

This point leads into a discussion of why racial hatred is treated differently, and furthermore whether any other defences are required, to be applicable across all forms of hatred. In particular, with the concerns currently present in the civil provision, including the lack of clarity as to what behaviour the provision is intended to be directed at, perhaps further exceptions would be beneficial.

The fact that other jurisdictions provide protections for freedom of expression for stirring up offences, aside from in regard to racial hatred, provides an impression that particular kinds of expression may not be protected in the context of racial hatred where it would be in other contexts. In the view of the UK Law Commission, the current racial hatred offence does not preclude against discussion of issues such as immigration, or other controversial topics subject to public debate.[173] However, this does not take account the idea of self-censorship and the chilling effect a lack of clarity can have on individual's freedom of expression, in terms of a fear of legal repercussions, which is also an issue present in New Zealand's proposals as they stand.

Without clarification, there is potential for politicians to be accused of hate speech when discussing legitimate issues of public concern, such as in the UK Conservative Party member was reported for hate speech for suggesting that employers should have to disclosure their number of foreign employees.[174] Immigration and other public policy decisions must be afforded the ability for robust debate, particularly when considering the democratic ideal justification for freedom of expression. This lead to the UK Law Commission suggesting a specific exception such as "*criticism or expressions of antipathy, dislike, ridicule, insult or abuse of particular countries or their governments*", which would prevent unjustified accusations of hate speech and clarify the scope of the racial hatred provision, which should also be a consideration for New Zealand.

The Canadian approach to exceptions should also be considered, where exceptions to s 319(2), the criminal offence of wilful promotion of hatred, apply across protected groups. Protected groups are deemed to be those based upon "colour, race, religion, national or ethnic origin, age, sex, sexual

[173] At [18.284].
[174] BBC News, "Joshua Silver reports Amber Rudd to police for hate speech" (12 January 2017) <https://www.bbc.co.uk/news/video_and_audio/headlines/38600100/joshua-silver-reports-amber-rudd-topolice-for-hate-speech>.

orientation, or mental or physical disability",[175] again, providing support for New Zealand widening protected groups. The exceptions, as set out in s 319(3), are as follows:

> *(a) if he establishes that the statements communicated were true;*
> *(b) if, in good faith, the person expressed or attempted to establish by an argument an opinion on a religious subject or an opinion based on a belief in a religious text;*
> *(c) if the statements were relevant to any subject of public interest, the discussion of which was for the public benefit, and if on reasonable grounds he believed them to be true; or*
> *(d) if, in good faith, he intended to point out, for the purpose of removal, matters producing or tending to produce feelings of hatred toward an identifiable group in Canada*

This further supports the notion that an exception for religious purposes should be considered in New Zealand's reforms, in order to ensure religious opinion and discussion are not stifled. Notably, the third exception could also be beneficial in order to ensure matters of public importance can still be freely debated, with particular regard to the proposal to extend protection to groups based upon political opinion. The inclusion of an exception of this nature in New Zealand's proposals could significantly help in clarifying the scope of the provisions, and allow for a wider protection of groups while still ensuring minimal and justified impairment on freedom of expression.

However, perhaps following the Scottish approach to exceptions would be a better option for New Zealand than using the Canadian ones. While the inclusion of a 'truth' defence is often supported on the basis that it is important to protect the 'marketplace of ideas' justification,[176] that operates on the basis of seeking truth, this does not necessarily protect the harm associated with hate speech. What may be 'true' at a certain point in time in a society may change, and thus the validity of the statement, yet the harmful effects of the statement will remain in relation to the targeted group.

Particularly in relation to civil hate speech legislation, courts in Canada have recognised that truth should not provide a shield in a human rights context and as truthful statements "can be interlaced with harmful ones or otherwise presented in a manner that would meet the definition of hate speech".[177] Even in regard to the criminal provision, where the truth defence is included, the Court in *Keegstra* questioned it's operation in both a practical and theoretical sense. With the standard of 'intent' being required in a criminal context, CJ Dickson stated that there is perhaps "no reason

[175] Canadian Criminal Code RSC 1985 C 46, s 318(4).
[176] Grant Huscroft and Paul Rishworth, above n 22, at 206.
[177] *Saskatchewan (Human Rights Commission) v Whatcott,* above n 78, at [140].

why the individual who intentionally employs such statements to achieve harmful ends must under the Charter be protected from criminal censure".[178]

UK jurisdictions do not have a truth defence, and based upon this and the discussion in Canadian courts I suggest that there is no strong basis for it to be included in a New Zealand context. The NZ Government should consider implementing a combination of the Scottish and UK and Wales approaches, with a broad freedom of expression provision for all characteristics and a more specific one for religious expression, in order to clearly protect and allow for discussion and criticism.

Following the recommendations of the UK Law Commission, I also suggest that a freedom of expression protection for race be considered, and that protections should apply across all groups, in order to avoid the potential ad hoc application, as discussed in relation to religion and race. The other defences included in the Canadian legislation would likely be covered by a freedom of expression provision, with both religious discussion and discussions of public interest being covered by this.

Conclusion

It is apparent through comparative analysis with other jurisdictions that exceptions for criminal hate speech/ incitement provisions are a common way to calibrate legislation in a way that gives adequate effect to freedom of expression. With the inclusion of 'freedom of expression provisions', there is increased support for extending the protected groups of hate speech provisions, as it will ensure sufficient clarity in terms of what the legislation is intended to target.

By ensuring debate and discussion are excluded, the hate speech provisions will more clearly capture conduct at the extreme end of the spectrum, which aligns with Edgeler's argument that broad extension of groups protected is a justified limitation on freedom of expression provided the provisions denote a high threshold with sufficient clarity. There is still however, an existing concern regarding a lack of evidence of a practical need to extend the legislation to all groups proposed, thus further evidence may need to be collected. However, other jurisdictions have indicated that in the interests of legal parity, the fact that prosecutions may be rarer in relation to

[178] *R v Keegstra*, above n 10, at 781.

certain groups should not preclude the introduction of legislation for behaviour that is to the same high threshold.

In terms of the civil provision, the exceptions discussed relate to criminal provisions, although it may be an option to have a freedom of expression exception in the civil arena also. As discussed in Chapter Two, the current proposal relating to the civil provision lacks clarity, in terms of how adding 'hatred' onto the current wording of 'hostility' and 'contempt' would be interpreted by the court and the extent to which this would lower the standard of the provision. A freedom of expression provision could therefore aid in addressing this, by ensuring the threshold can be lowered without infringing on legitimate discussion and debate.

Conclusion

Addressing hate speech in legislation is a complicated and controversial issue, with deeply held opinions present on both sides of the debate. New Zealand has been forced to deal with this issue following the 2019 terrorist attack on the Christchurch masjidain, and increasing beliefs that our current legislation is ineffective in addressing the harms of hate speech. However, hate speech legislation concurrently faces resistance on freedom of expression grounds, although as discussed, there is theoretical justification for hate speech legislation limiting freedom of expression to at least some degree based upon its underpinning principles. This has created a climate of disagreement on this issue, and therefore a careful balancing act must occur in the Government's recalibration of hate speech legislation.

This dissertation has looked to explore the Government's proposals objectively, looking first to the underlying theoretical debate, before turning to examine the legal implications of the proposed changes to the Human Rights Act 1993. Through comparative analysis with other jurisdictions and discussion of legal commentary, it is evident that aspects of the Government's proposals could be adapted in order to recalibrate the balance with freedom of expression in a more balanced way, thus there have been a number of recommendations made throughout this dissertation. However, it is recognised that the outcome of these proposals and resulting submissions may look different to what has been concluded throughout this dissertation, as the legislation will likely face significant debate and competing political motivations throughout its drafting stages.

Like most areas of the law, hate speech is about finding a balance between competing interests. However, this is exacerbated by the fact that hate is a complex human emotion that is difficult to address through legislation, as well as the current social climate where hate is increasingly being expressed through intolerance and violence. While most are familiar with the 'sticks and stones may break my bones' adage, it is becoming increasingly clear that words can also hurt, but less clear as to what point these words should be illegal. While it is perhaps unlikely that this question will ever receive a unanimous answer, it important that this question is carefully considered by New Zealand throughout this reform process, with a very careful balance needing to be struck.

Bibliography

A Cases

I New Zealand

Day v NZME Radio Ltd [2019] NZBSA 25.

Hansen v R [2007] NZSC 7.

King-Ansell v Police [1979] 2 NZLR 531.

Moonen v Film and Literature Board of Review [2000] 2 NZLR 9.

Morse v Police [2011] NZSC 45.

Wall v Fairfax New Zealand Ltd [2017] NZHRRT 17.

Wall v Fairfax New Zealand Ltd [2018] NZHC 104.

II Canada

Nealy v Johnston (1989) 10 CHRR D/6450.

R v Keegstra [1990] 3 SCR 697.

R v Oakes [1986] 1 SCR 103.

Sakatchewan (Human Rights Commission) v Whatcott [2013] 1 SCR 467.

Taylor v Canadian Human Rights Commission [1990] 3 SCR 892 at 928.

III United Kingdom

Mandla v Dowell Lee [1983] 2 AC 548 (HL).

Miller v College of Policing [2020] EWHC 225.

IV United States

Abrams v United States 250 U.S. 6'6 R 620 (919).

Schenck v. United States 249 US 49 (1918).

B Legislation

I New Zealand

Broadcasting Act 1989.

Human Rights Act 1993.

New Zealand Bill of Rights Act 1990.

II Canada

Canadian Criminal Code RSC 1985 C 46.

Canadian Human Rights Act, RSC 1985, c H-6.

The Canadian Charter of Rights and Freedoms, pt 1 of the Constitution Act 1982, being sch B to the Canada Act 1982 (UK).

III United Kingdom

Public Order Act 1986 (UK).

Hate Crime and Public Order (Scotland) Act 2021.

C Books and Chapters in Books

Christine Sanders (ed) Human Rights Law (online looseleaf ed, Thomson Reuters) at [HR61.05].

Grant Huscroft and Paul Rishworth Rights and Freedoms: The New Zealand Bill of Rights Act 1990 and the Human Rights Act 1993 (Wellington, Brookers, 1995).

Grant Huscroft "Defamation, Racial Disharmony, and Freedom of Expression" in Grant Huscroft and Paul Rishworth (eds) Rights and Freedoms: The New Zealand Bill of Rights Act 1990 and the Human Rights Act 1993 (Brookers, Wellington, 1995) 171.

Jeremy Waldron The Harm in Hate Speech (Harvard University Press, Cambridge, 2012).

John Stuart Mill On Liberty (London: Longman, Roberts and Green, 1869).

Kent Greenawalt Fighting Words: Individuals, Communities, and Liberties of Speech (Princeton University Press, Princeton (NJ), 1995).

Paul Rishworth, et al. The New Zealand Bill of Rights (Oxford: Oxford University Press, 2003) at 309.

Paul Rishworth The Right to Freedom of Expression (A research paper written for the New Zealand Human Rights Commission) in Christine Sanders (ed) Human Rights Law (online looseleaf ed, Thomson Reuters) at [HR61.05].

Katharine Gelber and Adrienne Stone Hate speech and freedom of speech in Australia (Federation Press, Annadale (NSW), 2007) at 67.

Stavros Assimakopoulos, Fabienne H. Baider and Sharon Millar Online Hate Speech in the European Union A Discourse-Analytic Perspective (1st ed, Springer Nature, 2017).

D Journal Articles

Dominic McGoldrick and Therese O'Donnell, T "Hate-speech laws: Consistency with national and international human rights law" (1998) 18(4) Legal Studies 453-485.

Eugene Volokh "In Defense of the Marketplace of Ideas / Search For Truth as a Theory of Free Speech Protection" (2011) 97(3) Virginia Law Review 595-603.

Jeffrey Howard, "Should we ban dangerous speech?" (2018) 32 British Academy Review 19.

Katharine Gelber and Luke McNamara "Evidencing the harms of hate speech" (2016) 22 Social Identities Journal for the Study of Race, Nation and Culture at 324-341.

Kathleen Mahoney "Hate Vilification Legislation and Freedom of Expression: Where Is the Balance?" [1994] 21 AUJlHRights.

Lauren E Scharfstein "The Hate Speech Debate: The Supreme Court, the Federal Government, and the Need for Civil Hate Speech Provisions" (2019) 19 Asper Review of International Business and Trade Law 375 at 382.

Mari J. Matsuda "Public response to racist speech: Considering the victim's story" (1989) 87 Michigan Law Review 2320 at 2327.

Matthew L Williams and others "Hate in the Machine: Anti-Black and Anti-Muslim Social Media Posts as Predictors of Offline Racially and Religiously Aggravated Crime" (2020) 60 British Journal of Criminology 93–117.

Morton J. Horwitz "Rights" (1998) 23 HARV. C.R.-C.L. L. REV. 393-406.

E Government Materials

Ministry of Justice Interim Impact Summary: Public discussion document – Proposed changes to the incitement provisions in the Human Rights Act 1993 (18 June 2021).

Ministry of Justice Proposals against incitement of hatred and discrimination (25 June 2021) at 30.

Ministry of Justice "Proposals against incitement of hatred and discrimination in Aotearoa New Zealand: Public submissions are now closed" Justice.govt.nz <https://www.justice.govt.nz/justice-sector-policy/key-initiatives/proposals-against-incitement/>.

International Agreements

International Convention on the Elimination of All Forms of Racial Discrimination 660 UNTS 195 (opened for signature 21 December 1965, entered into force 4 January 1969).

International Covenant on Civil and Political Rights 999 UNTS 171 (entered into force on 23 March 1976).

F Reports

House of Commons (Canada) TAKING ACTION TO END ONLINE HATE: Report of the Standing Committee on Justice and Human Rights (June 2019).

Ko tō tātou kāinga tēnei: Royal Commission of Inquiry into the terrorist attack on Christchurch masjidain on 15 March 2019, (26 November 2020).

Law Commission (UK) Hate Crime Laws: A Consultation Paper (23 September 2020).

Law Commission (UK) Hate Crime: Should the Current Offences be Extended? (May 2014).

Lord Bracadale Independent Review of Hate Crime Legislation in Scotland Final Report (31 May 2018).

Netsafe Online Hate Speech: A survey on personal experiences and exposure among adult New Zealanders (November 2018).

New Zealand Human Rights Commission Kōrero Whakamauāhara: Hate Speech An overview of the current legal framework (December 2019)

New Zealand Human Rights Commission Human Rights in New Zealand 2010 – Nga Tika Tangata O Aotearoa 2010 (10 December 2010).

The Scottish Parliament Hate Crime and Public Order (Scotland) Bill Policy Memorandum (April 2020).

The Scottish Government, One Scotland: Hate has no home here – Consultation on amending Scottish hate crime legislation Analysis of responses: Final Report (June 2019) at 4.

G Dissertations

Elizabeth MacPherson, "Regulating Hate Speech in New Zealand" (Wellington 2003) Victoria University of Wellington Faculty of Law.

H United Nations Materials

Frank La Rue Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression UN Doc A/67/357 (7 September 2012)

United Nations General Assembly Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred UNDoc A/HRC/22/17/Add.4 (11 January 2013)

United Nations UN Strategy and Plan of Action on Hate Speech (May 2019).

I Internet Resources

BBC News, "Joshua Silver reports Amber Rudd to police for hate speech" (12 January 2017) <https://www.bbc.co.uk/news/video_and_audio/headlines/38600100/joshua-silver-reports-amber-rudd-topolice-for-hate-speech>.

Charlie Dreaver "Hate speech laws: ACT, National say proposals a divisive overreach" (25 June 2021) RNZ <https://www.rnz.co.nz/news/political/445548/hate-speech-laws-act-national-say-proposals-a-divisive-overreach>.

David Seymour "PETITION: Stop Labour's Hate Speech Laws" (2021) ACT New Zealand <https://freespeech.act.org.nz/>.

Eddie Clark "Hate speech law must be calibrated carefully" (15 July 2021) Star News <https://www.odt.co.nz/star-news/star-opinion/hate-speech-law-must-be-calibrated-carefully>.

Eddie Clark "NZ's hate speech proposals need more detail and wider debate before they become law" (27 April 2021) The Conversation <https://theconversation.com/nzs-hate-speech-proposals-need-more-detail-and-wider-debate-before-they-become-law-159320>.

Graeme Edgeler "The Government's Proposed Decriminalisation of Racist Hate Speech" (28 June 2021) Legal Beagle <https://publicaddress.net/legalbeagle/the-governments-proposed-decriminalisation/>.

Graeme Edgeler (11 July 2021) "Political opinion and the proper scope of hate speech laws; a post in honour of John Campbell" Legal Beagle <https://publicaddress.net/legalbeagle/political-opinion-and-the-proper-scope-of/>

New Zealand National Party (@NZNationalParty) " National will reverse any attempts Jacinda Ardern's Government makes to criminalise speech beyond the threshold of 'inciting violence.'" <https://twitter.com/nznationalparty/status/1410112520514277377>.

Stephen Price "Hating On Hate Speech Laws" (29 June 2021) Media Law Journal <https://www.medialawjournal.co.nz/?p=733>.

Stephen Price "Hate Speech: a question" (5 July 2021) Media Law Journal <https://www.medialawjournal.co.nz/?p=744>.

Appendix 1

Human Rights Act 1993

61 Racial disharmony

(1) It shall be unlawful for any person—

> (a) to publish or distribute written matter which is threatening, abusive, or insulting, or to broadcast by means of radio or television or other electronic communication words which are threatening, abusive, or insulting; or
>
> (b) to use in any public place as defined in section 2(1) of the Summary Offences Act 1981, or within the hearing of persons in any such public place, or at any meeting to which the public are invited or have access, words which are threatening, abusive, or insulting; or
>
> (c) to use in any place words which are threatening, abusive, or insulting if the person using the words knew or ought to have known that the words were reasonably likely to be published in a newspaper, magazine, or periodical or broadcast by means of radio or television,—

being matter or words likely to excite hostility against or bring into contempt any group of persons in or who may be coming to New Zealand on the ground of the colour, race, or ethnic or national origins of that group of persons.

131 Inciting racial disharmony

(1) Every person commits an offence and is liable on conviction to imprisonment for a term not exceeding 3 months or to a fine not exceeding $7,000 who, with intent to excite hostility or ill-will against, or bring into contempt or ridicule, any group of persons in New Zealand on the ground of the colour, race, or ethnic or national origins of that group of persons,—

> (a) publishes or distributes written matter which is threatening, abusive, or insulting, or broadcasts by means of radio or television words which are threatening, abusive, or insulting; or
>
> (b) uses in any public place (as defined in section 2(1) of the Summary Offences Act 1981), or within the hearing of persons in any such public place, or at any meeting to which the public are invited or have access, words which are threatening, abusive, or insulting,—

being matter or words likely to excite hostility or ill-will against, or bring into contempt or ridicule, any such group of persons in New Zealand on the ground of the colour, race, or ethnic or national origins of that group of persons.