NZ Human Rights.
Human Rights Commission
Te Kāhui Tika Tangata

Kōrero Whakamauāhara: Hate Speech

An overview of the current legal framework

Human Rights Commission
Te Kāhui Tika Tangata

December 2019

"When expressions of contempt for one group become normative, it is virtually inevitable that similar hatred will be directed at other groups. Like a fire set by an arsonist, passionate hatred and conspiratorial worldviews reach well beyond their intended target."

Deborah E. Lipstadt, Antisemitism: Here and Now, 2019.

ISSN: 978-0-478-35644-1 (Print)
ISSN: 978-0-478-35645-8 (Online)

Published December 2019
Wellington, Aotearoa, New Zealand

Foreword

On 15 March 2019, fifty-one people were killed at their place of worship. The Christchurch mosque massacres shook the country's foundations and generated deep soul-searching.

People asked, what was the source of this hate? Have we been doing enough to tackle Islamophobia, racism and antisemitism? As a multicultural society, based on Te Tiriti o Waitangi, how can we defeat these and other forms of hate, for example, against disabled people, women and the rainbow community? How can we promote an inclusive society that encompasses all individuals and communities, including those who feel silenced and forgotten? What are the core values that should guide our collective response to 15 March?

As these questions imply, the appropriate collective response to the massacres will include local, national and international initiatives. It will encompass a variety of policies, programmes and practices. The legislation prohibiting semi-automatic weapons, enacted in April this year, demonstrates that law has a role to play.

In short, the response to 15 March will be complex and multifaceted. This report addresses only one corner of this large canvass: hate speech, also known as harmful speech.

Are New Zealand's existing laws on hate speech fit for purpose? Does the current legal framework strike the correct balance between freedom of expression and other human rights, such as non-discrimination and equality? Do the laws discharge the government's fundamental obligation to ensure that all individuals and communities are safe and secure from avoidable harm?

All responsible countries wrestle with these challenging questions. For example, in a leading US Supreme Court case, Justice Frank Murphy outlined when speech may be curtailed, including in relation to the "lewd and obscene, the profane, the libellous and the insulting or 'fighting' words – those which, by their very utterance, inflict injury or tend to incite an immediate breach of the peace."[1] As the present report demonstrates, the United Nations has much to say on hate speech. In 2004 a parliamentary committee reviewed New Zealand's hate speech laws but regrettably does not appear to have published a report.

This paper aims to provide an accessible introduction to hate or harmful speech in national and international law. It serves as a resource i.e. it need not be read from beginning to end. Some readers may prefer to read Part I which includes the reasons for and against the regulation of hate speech. Other readers may prefer to focus on Part II which introduces the relevant international human rights law. While others may wish to look at Part III on New Zealand's current hate speech laws, or Part IV on the approach of some other countries, including Australia and Canada, to hate speech.

Importantly, the report aims to be neither comprehensive nor an account of the position of the Human Rights Commission in relation to hate or harmful speech. Instead, the publication is another example of what the Commission has consistently tried to do since 15 March: provide a modest contribution to complex issues with a view to engendering well-informed, inclusive and respectful discussion.

Aotearoa New Zealand will not be defined by 15 March, it will be defined by our collective long-term response to the catastrophe.

Paul Hunt

Chief Human Rights Commissioner

[1] *Chaplinsky v New Hampshire* 315 US 568 (1942) at 571-572. Justice Felix Frankfurter cited Justice Frank Murphy in *Beauharnais v Illinois* 343 US 250 (1952) at 256.

Contents

Introduction

In a free society, the central question arising from any debate about hate speech and the law is where to draw the line. How do we determine the boundary between expression that is hateful or morally objectionable, though lawful, and that which is of such a harmful nature that it should be subject to legal sanctions?

The term hate speech can be misleading because it is often used loosely and pejoratively to imply a moral breach and is directed at the speech or expression itself. However, hate speech laws are not intended to protect people from offence or to supress ideas. They are targeted at the effect that the expression has on the minds of third parties.

Hate speech laws in a number of countries were drafted in the post-World War II era, where the focus was on supressing racially motivated hate speech rather than speech directed at other types of personal characteristics. However, in recent years discussion about the type of hate speech that should be prohibited has widened. This has led to changes in the law in countries such as England, Canada and Australia.

In New Zealand, there have been similar public discussions about what kind of speech, if any, should be subject to regulation.

Discussions of this nature require consideration of both the right to freedom of expression (which includes the right to freedom of opinion), and the right to equality and non-discrimination. These two rights are set out in various international human rights treaties and in the New Zealand Bill of Rights Act 1990. The right to freedom of expression is a qualified right meaning that it can be restricted in certain circumstances.

This paper aims to provide readers with information about the legal regulation of hate speech under international and domestic laws. The first part of this paper provides an overview of what hate speech laws are and provides arguments for and against them. The second part sets out the international laws that regulate speech. The third part of the paper provides information on the current regulation of hate speech in New Zealand. Lastly, it looks at hate speech laws in similar jurisdictions to New Zealand.

Part I: Hate speech

What is hate speech?

The term hate speech is often used in an emotive sense to describe speech or expression that members of society deem as morally wrong. While speech may be morally reprehensible, it is not necessarily unlawful. This paper will focus on what international law says about the regulation of hate speech and how some comparative countries regulate hate speech.

No definition of hate speech exists under international law and definitions under national laws vary. In general, hate speech laws usually cover not only spoken words, but words or images printed, published or posted on the internet.[2] This is particularly important today because material that is printed or published online has the potential to reach large audiences.

Discussions about hate speech often focus on the ideas being expressed, rather than the effect of the expression.[3] However, the aim of hate speech laws is not usually to supress expressions of hate in themselves, but to punish the incitement of hate in other individuals. As highlighted by the Human Rights Review Tribunal in New Zealand's most recent case under the civil hate speech provision found in section 61 of the Human Rights Act 1993:[4]

> *[T]he harm at which the legislation is directed is not the expression itself, but the presumed effect of that expression on the minds of third parties.*

The leading Canadian Supreme Court case of *Saskatchewan Human Rights Commission v Whatcott* also made this point. The Court was tasked with determining whether the hate speech provision under the Saskatchewan Human Rights Code was constitutional in light of the right to freedom of expression. In coming to its decision, the Court identified that one of the challenges of the application of hate speech laws is the "mistaken propensity to focus on the ideas being expressed, rather than on the *effect* of the expression."[5]

It is notable that in the United Kingdom hate speech laws are referred to as "stirring up" offences, which provides a good description of the intended purpose of such legislation.

Why regulate hate speech?

Scotland's 2017 independent review of hate crime legislation found that the merits of having "stirring up" offences included recognition of moral wrongfulness, harm, the seriousness of the offence as well as their symbolic nature.[6]

In concluding that Scotland's stirring up offences should be extended, Lord Bracalade concluded that "*the harm caused by stirring up of hatred offences can be particularly severe and it is an important consideration pointing towards the extension of such offences.*"[7]

---

[2] Jeremy Waldron, *The Harm in Hate Speech* (Harvard University Press, Cambridge, 2012) at 37.
[3] At 492.
[4] *Wall v Fairfax* [2017] NZHRRT 17 at [191.3].
[5] *Saskatchewan Human Rights Commission v Whatcott* [2013] SCC 11 at 492 [31]. See also: *The distinction between the expression of repugnant ideas and expression which exposes groups to hatred is crucial to understanding the proper application of hate speech prohibitions. Hate speech legislation is not aimed at discouraging repugnant or offensive ideas. It does not, for example, prohibit expression which debates the merits of reducing the rights of vulnerable groups in society. It only restricts the use of expression exposing them to hatred as a part of that debate. It does not target the ideas, but their mode of expression in public and the effect that this mode of expression may have at* [51]; *and… An assessment of whether expression exposes a protected group to hatred must therefore include an evaluation of the likely effects of the expression on its audience. Would a reasonable person consider that the expression vilifying a protected group has the potential to lead to discrimination and other harmful effects? This assessment will depend largely on the context and circumstances of each case at* [52].
[6] Lord Bracadale, *Independent Review of Hate Crime Legislation in Scotland – Final Report* (Scottish Government, Edinburgh, May 2018) at iv and 9.
[7] At 58.

The Canadian Supreme Court case of *R v Keegstra*, which examined the constitutionality of Canada's federal hate crime offence, identified the impact on society at large and the discord that hate speech can create between groups.[8]

New York University Professor Jeremy Waldron, author of *The Harm in Hate Speech*, argues that regulating hate speech protects the public good of dignity-based assurance:[9]

> *Dignity . . . is precisely what hate speech laws are designed to protect-not dignity in the sense of any particular level of honor or esteem (or self-esteem), but dignity in the sense of a person's basic entitlement to be regarded as a member of society in good standing, as someone whose membership of a minority group does not disqualify him or her from ordinary social interaction. That is what hate speech attacks, and that is what laws supressing hate speech aim to protect.*

He argues that offence, however deeply felt, is not a proper object of legislative concern but that:

> *Dignity and the assurance that comes with it are public goods constituted by what thousands or millions of individuals say and do. Our society is heavily invested in the provision of those goods. The point of hate speech is to detract from that provision – to undermine it and establish rival goods that indicate (to fellow racists, to members of vulnerable groups, and to society generally) that the position of some minority or other is by no means as secure as the rest of the world would like to affirm. The point of hate speech restrictions…is to protect the first set of public goods from being undermined in this way.*
>
> *They are set up to vindicate public order, not just by pre-empting violence, but by upholding against attack a shared sense of the basic elements of each person's status, dignity, and reputation as a citizen or member of society in good standing—particularly against attacks predicated upon the characteristics of some particular social group.*

Hate speech may lead to violence or public disorder because it encourages the commission of hate crimes against members of the targeted groups.[10] Hate speech may also lead to the promotion of negative stereotypes about the targeted group, resulting in an increase in discriminatory treatment of members of that group in society. While it is difficult to establish empirical evidence, it is likely that the cumulative effect over time of hate speech targeted at a particular group will be to spread and reinforce negative social attitudes and escalate discrimination towards members of the targeted group.

The Human Rights Review Tribunal in *Wall v Fairfax* referred to the impact of hate speech:[11]

> *While the right to freedom of expression is one of the most essential elements in a democratic society we accept the principle that expression that advocates racial disharmony or hatred against a group of persons on the basis of their immutable*

[8] *R v Keegstra* [1990] 3 SCR 697: Essentially, there are two sorts of injury caused by hate propaganda. First, there is harm done to members of the target group. It is indisputable that the emotional damage caused by words may be of grave psychological and social consequence … a response of humiliation and degradation from an individual targeted by hate propaganda is to be expected. A person's sense of human dignity and belonging to the community at large is closely linked to the concern and respect accorded the groups to which he or she belongs. The derision, hostility and abuse encouraged by hate propaganda … have a severely negative impact on the individual's sense of self-worth and acceptance. This impact may cause target group members to take drastic measures in reaction, perhaps avoiding activities which bring them into contact with non-group members or adopting attitudes and postures directed towards blending in with the majority. Such consequences bear heavily in a nation that prides itself on tolerance and the fostering of human dignity through, among other things, respect for the many racial, religious and cultural groups in our society. A second harmful effect of hate propaganda … is its impact on society at large. It is thus not inconceivable that the active dissemination of hate propaganda can attract individuals to its cause, and in the process create serious discord between various cultural groups in society.

[9] Jeremy Waldron, The Harm in Hate Speech (Harvard University Press, Cambridge, 2012) at 105.

[10] At 73 cites commentators that have utilised historical examples to draw a link between hate speech and genocide.

[11] Wall v Fairfax [2017] NZHRRT 17 at [186]; the Tribunal referred to the European Court of Human Rights in Vona v Hungary [2013] ECHR 653 at [57]: In the Court's view, the State is also entitled to take preventive measures to protect democracy vis-à-vis such non-party entities if a sufficiently imminent prejudice to the rights of others threatens to undermine the fundamental values on the basis of which a democratic society exists and functions. One such value is the coexistence of members of society free from racial segregation, without which a democratic society is inconceivable.

> *characteristics is harmful to the achievement of the values of a democratic society which respects (inter alia) human dignity, equality and fundamental freedoms including the right to be free from discrimination.*

Moana Jackson has pointed out that repetitive hatred cannot breed compassion and it carries a cost to those affected.[12] Jackson points to the history of colonisation where free speech has been used to excuse and maintain privilege and the devaluing of indigenous values, language and rights.

Arguments against regulating hate speech

The arguments against regulating hate speech centre on freedom of expression. Freedom of expression has become one of the world's most widely recognised rights.[13] For some concerned with civil liberties, there is a strong belief that everyone must be able to express their opinion regardless of how worthless or odious it may be thought to be.[14] Classical liberal discourse provides numerous rationales for freedom of speech and consequently against regulating hate speech. These include that it is essential for democracy; guarantees the marketplace of ideas; and promotes individual autonomy. These and other common reasons put forward for not regulating hate speech are outlined below.

The first key argument is based on the recognition that freedom of speech is essential to democracy.[15] A democracy is a society in which the citizens debate and decide the laws for themselves. To be a true democracy, then, citizens must be free to discuss any idea, no matter how repugnant it may be.[16] This enables voters to be better informed and allows state officials to be held accountable for their actions. Freedom of expression is also essential for a representative government. By facilitating public discussion on controversial issues, the government can recognise and combat social problems more effectively and those in Parliament can better represent their constituents.[17] Some have noted however that this rationale implies hate speech can be limited when it does not influence the democratic process.[18]

The second concerns the concept of the "marketplace of ideas" which recognises the advancement of knowledge and the discovery of truth as a fundamental good.[19] It treats people as adults which requires letting them hear bad ideas, so that they can make up their minds for themselves.[20] As stated by the political philosopher John Stuart Mill in *On Liberty*:[21]

> *But the peculiar evil of silencing an opinion is that it is robbing the human race; posterity as well as the existing generation; those who dissent from the opinion, still more than those who hold it. If the opinion is right, they are deprived of the opportunity of exchanging error for truth; if wrong, they lose what is almost as great a benefit, the clearer perception and livelier impression of truth, produced by its collision with error.*

This theory was the basis for the dissenting opinion in the Supreme Court of the United States case of *Abrams v US*.[22] Legal scholar, Bollinger, states that freedom of expression promotes the "right" attitudes of tolerance among the audience and performs a self-reformation

[12] Moana Jackson, "Rethinking free speech" (19 August 2018) E-Tangata, <https://e-tangata.co.nz/comment-and-analysis/moana-jackson-rethinking-free-speech/>.
[13] Dominic McGoldrick and Terese O'Donnell "Hate-Speech Laws: Consistency with National and International Human Rights Law" (1998) 18 Legal Studies (Society of Public Teachers of Law) at 454.
[14] Grant Huscroft and Paul Rishworth, Rights and Freedoms (Wellington, Brookers, 1995) at 192.
[15] Rex Ahdar, "The Right to Protection of Religious Freedom" (2008) OLR 363.
[16] Jeffrey Howard "Should we ban dangerous speech?" (2018) 32 British Academy Review 19 at 20.
[17] Hannah Musgrave "What Makes Race So Special? Should hate speech provisions under the Human Rights Act 1993 be extended to cover target groups other than race?" (2009) OYLR at 5.
[18] At 5.
[19] *Abrams v US* (1919) 250 US 616 at 630.
[20] Jeffrey Howard "Should we ban dangerous speech?" (2018) 32 British Academy Review 19 at 20.
[21] John Stuart Mill, *On Liberty* (London: Longman, Roberts and Green, 1869) cited in John Gray and G.W. Smith, J.S. Mill *On Liberty: In Focus* (London and New York: Routledge, 1991) 17. See also reference in Geoffrey Palmer, "The purposes of protecting speech" [2018] NZLJ 227 at 227.
[22] *Abrams v US*, above n 19, at 616.

function for the general community.[23] In other words, the right way to counter hateful ideas is to allow them to be tested in the field of open debate, where they will eventually be refuted conclusively and wither away. As stated by Lester and Bindman:[24]

> *Democracy stands (and some would say, may fall) on the conviction that unpopular ideas would be freely expressed, and that, if they are false or evil, they will ultimately be defeated, not by censorship or prosecution, but by public education and debate.*

However, it has been argued that if hate speech leads to its targets retracting their opinions from public consumption then Mill's marketplace of ideas is compromised.[25]

The third key argument against regulating hate speech is that part of respecting people's autonomy is letting them express who they are and what they think, even if we vehemently disapprove.[26] This has also been expressed as the promotion of self-development and personal fulfilment. The personal autonomy argument allows individuals to weigh up competing ideas and develop a sense of self when preferring a view over other views. However, the counterview is that by creating a climate of discrimination, hate speech undermines self-worth and causes targets to encompass a subordinate status which some argue runs counter to the self-development rational.[27]

Other arguments against regulating free speech include the view that laws restricting speech are counter-productive, allowing racists to use the courts to achieve far more publicity for their racist viewpoints than otherwise may have been achieved.[28] Related to this is the view that it allows racists to enhance their personal status within their community, acting as a sort of martyr to the cause.[29] It is also argued that hate speech laws antagonise citizens, forcing them underground and making them more dangerous.[30] Instead, by permitting the expression of dangerous ideas, we are better able to defuse their danger by arguing against them.[31] A further argument against regulation of hate speech is that it is dangerous to give the state the power to restrict dangerous speech, since it is likely to misuse that power (either by making mistakes about what is genuinely dangerous, or by abusing the power for political purposes).[32]

Hate crime compared with hate speech

While hate speech and hate crime are connected, and can overlap, they are distinct concepts. As noted above, hate speech addresses words or expression that encourage others to hate a group based on a protected characteristic. In most hate speech legislation, hate is primarily relevant, not as the motive for the words spoken but as a possible effect of the form of speech.[33] The words, whether or not spoken hatefully by the perpetrator, result in the incitement of hostility or hatred against others.

Hate crimes, on the other hand, involve the commission of an offence, for example an assault against a person or damage to property, which is accompanied by the motive of hatred against a protected group. It is the motive or demonstration of hostility that marks it out as a hate crime.

New Zealand does not have specific hate crime offences. However, hateful motivations are considered an aggravating factor in sentencing under the

[23] Lee Bollinger, The Tolerant Society: Freedom of Speech and Extremist Speech in America (New York, Oxford University Press, 1986).
[24] Anthony Lester and Geoffrey Bindman, *Race and Law* (London, Penguin, 1972) at 358.
[25] Christopher Jones, "Rocks can Turn to Sand and Wash Away But Words Last Forever: A Policy Recommendation for New Zealand's Vilification Legislation", (Hamilton, University of Waikato, 2002) at 13.
[26] Jeffrey Howard, "Should we ban dangerous speech?" (2018) 32 British Academy Review 19 at 20.
[27] Elizabeth MacPherson, "Regulating Hate Speech in New Zealand" (Wellington. Victoria University of Wellington Faculty of Law, 2003) 20.
[28] Grant Huscroft and Paul Rishworth, *Rights and Freedoms* (Wellington, Brookers, 1995) at 193.
[29] At 193.
[30] Jeffrey Howard, "Should we ban dangerous speech?" (2018) 32 British Academy Review 19 at 20; Grant Huscroft and Paul Rishworth, *Rights and Freedoms* (Wellington, Brookers, 1995) at 193.
[31] Jeffrey Howard, "Should we ban dangerous speech?" (2018) 32 British Academy Review 19 at 20.
[32] At 20.
[33] Jeremy Waldron, *The Harm in Hate Speech* (Harvard University Press, Cambridge, 2012) at 35.

Sentencing Act 2002. Section 9(1) of the Act requires the courts to take into account the following factors in sentencing a convicted offender:

> *(h) that the offender committed the offence partly or wholly because of hostility towards a group of persons who have an enduring common characteristic such as race, colour, nationality, religion, gender identity, sexual orientation, age, or disability; and*
>
> *(i) the hostility is because of the common characteristic; and*
>
> *(ii) the offender believed that the victim has that characteristic.*

The list of protected characteristics under section 9(1)(h) is relatively wide ranging, covering race, colour, nationality, religion, gender identity, sexual orientation, age and disability. Approaches to hate crime laws and associated data collection have been the subject of debate in New Zealand.[34] There have also been repeated calls for better collection of data about crimes motivated by hate to ensure that there is visibility and understanding about the prevalence and nature of such events and to allow better targeting of resources in response.[35]

Section 131 of the Human Rights Act also makes certain types of hate speech a criminal offence. The threshold is high and consent must be obtained from the Attorney-General before any prosecution can be commenced for this offence. Section 61 of the Human Rights Act contains a similar civil prohibition on certain types of hate speech. These provisions are discussed in more detail later in this paper.

Balancing and limiting rights

When considering the issue of hate speech, one of the challenges is reconciling the need to protect and promote the right to freedom of opinion and expression, on the one hand, and to combat intolerance, discrimination and incitement to hatred, on the other.[36] While hate speech laws may challenge the right to freedom of expression, they are not designed to hinder words or debate.[37] In the context of this discussion it is important to be clear about the distinction between *explicit limitations* on a right and the *balancing* of different rights. Freedom of speech, like most human rights, is not absolute and it is subject to certain restrictions or limitations under international and national law. The challenge is to precisely define and apply these lawful limitations. In contrast, balance is a key feature or attribute that runs through human rights law and practice. Balances have to be struck between human rights. For example, the right to privacy has to be balanced with the right to information. Equally, freedom of speech has to be balanced with other human rights such as the rights to security and non-discrimination.

The right to freedom of expression is one of the most valued and well known of all human rights. It is protected in article 19 of the Universal Declaration of Human Rights (UDHR) and article 19 of the International Covenant for Civil and Political Rights 1966 (ICCPR). It requires States to guarantee to all people the freedom to seek, receive or impart information or ideas of any kind, either orally, in writing or in print, in the form of art, or through any other media of a person's choice.

The right to freedom of expression is reflected in New Zealand law under the New Zealand Bill of Rights Act 1990 (BORA). Section 14 reads:

> *Everyone has the right to freedom of expression, including the freedom to seek, receive, and impart information and opinions of any kind in any form.*

[34] See for example, Sam Sachdeva, "Government downplays police suggestion of new hate crime legislation after Muslim abuse in Huntly" Stuff (15 February 2017); Joe Hingham "Invisible violence: Why we need a hate crime law in New Zealand" The Spinoff (16 November 2017); Sara Vui-Talitu "Muslim Kiwi says hate crime legislation must change" RNZ (17 June 2019).

[35] See for example, "It Happened Here -reports of race and religious hate crime in New Zealand 2004-2012" (June 2019). https://www.hrc.co.nz/files/1515/6047/9685/It_Happened_Here_Reports_of_race_and_religious_hate_crime_in_New_Zealand_2004-2012.pdf

[36] United Nations General Assembly, 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 [3].

[37] Although it should be noted that one of the criticisms of hate speech laws generally is that they can be used by States as a tool to deter or punish legitimate criticism of government action, particularly in non-democratic States or fragile democracies.

Freedom of expression is one of the essential foundations of a democratic society because it guarantees the right of every person to exchange information, debate ideas and express opinions.[38] It encompasses the expression of opinions and ideas that others may find deeply offensive, and it may encompass discriminatory expression.[39]

Despite its fundamental importance, the right to freedom of expression is not absolute and it can be restricted in some circumstances. In the context of hate speech, international law provides for the type of restrictions that may be appropriate. These are outlined in Part II.

In New Zealand, the BORA recognises that the right to freedom of expression can be subject to "reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society."[40] The test to be applied under section 5 was set by the Supreme Court in its landmark judgment in *Hansen v R*, which drew upon the decision of the Canadian Supreme Court in *R v Oakes*. The components of the Hansen test, set by Tipping J, are as follows:[41]

(a) Does the limiting measure serve a purpose sufficiently important to justify curtailment of the right or freedom?

(b) (i) Is the limiting measure rationally connected with its purpose?

(ii) Does the limiting measure impair the right or freedom no more than is reasonably necessary for sufficient achievement of its purpose?

(iii) Is the limit in due proportion to the importance of the objective?

The right to freedom of expression should not be aimed at the violation of any of the rights and freedoms of others, including the right to equality and non-discrimination.[42] International human rights law guarantees equality and non-discrimination for all people.[43] The principle of non-discrimination has three elements:[44]

- any distinction, exclusion, restriction or preference against a person;
- based on a protected characteristic recognised under international human rights law;
- which has the purpose or effect of nullifying or impairing the recognition, enjoyment or exercise, on an equal footing, of human rights and fundamental freedoms in the political, economic, social, cultural or any other field of public life.

[38] *Handyside v UK* [1976] 1 ECHR 5 at [49] as cited by the New Zealand Court of Appeal in *Living Word Distributors Ltd v Human Rights Action Group (Wellington)* [2000] 3 NZLR 570 at [45]: *Freedom of expression constitutes one of the essential foundations of a [democratic] society … it is applicable not only to 'information' or 'ideas' that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population. Such are the demands of that pluralism, tolerance and broadmindedness without which there is no 'democratic society'.*

[39] Human Rights Committee, General Comment No. 34: Article 19: Freedom of option and expression, UN Doc CCPR/C/GC/34, (12 September 2011) at [11].

[40] *Wall v Fairfax* [2017] NZHRRT 17 at [171]: *The overarching conclusion to be drawn from ICERD and the ICCPR is that neither the right to be free from hate speech nor the right to freedom of expression is an absolute right. The "due regard" clause in article 4 of ICERD and the qualifications in the third paragraph of article 19 of the ICCPR are explicit in this regard. At treaty level, conflict between these rights is to be resolved by the principles of legality, proportionality and necessity; and [172]: But when rights conflict in New Zealand domestic law, such conflict must be resolved within the analytical framework prescribed by domestic law. That framework is different to the international one not least because whereas article 4 of ICERD and article 19 of the ICCPR contain their own limitation clauses, there is none in s 14 of the Bill of Rights. The limitations are to be found elsewhere, that is in ss 4, 5 and 6 of the Bill of Rights. It is these provisions which, read together, provide the domestic framework of analysis. See further the discussion in Butler and Butler op cit [6.11.20].*

[41] *Hansen v R* [2007] NZSC 7 at [104].

[42] United Nations General Assembly, 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 [36].

[43] See article 1 of the UDHR and articles 2(1) and 26 of the ICCPR. At the regional level, for example, articles 2 and 9 of the African Charter on Human and Peoples' Rights; articles 1(1) and 24 of the American Convention of Human Rights; and, article 14 and Protocol 12 to the European Convention of Human Rights.

[44] Human Rights Committee, General Comment No. 18: Non-discrimination, XXXVII, (10 November 1989) at [6].

The right to equality and non-discrimination is affirmed under New Zealand law. Section 19(1) of the NZBORA reads:

> *Everyone has the right to freedom from discrimination on the grounds of discrimination in the Human Rights Act 1993.*

The right to freedom of expression can also be limited to promote other values that are considered to be of greater societal importance. For example, certain media, such as films, magazines or books may be censored or their distribution restricted under the Films, Videos, and Publications Classifications Act 1993; the Crimes Act sanctions threatening or offensive statements; and defamation laws exist.

The rise of online hate speech

The rapid increase in the use of social media has contributed to the prevalence of online hate speech. Hateful comments can be easily shared across a large audience and can be made with anonymity and invisibility. Recently the Special Rapporteur on the promotion and protection of the freedom of opinion and expression (Special Rapporteur), explained the harm of online hate speech in the following way:[45]

> *Online hate speech, […] can result in deleterious outcomes. When the phrase is abused, it can provide ill-intentioned States with a tool to punish and restrict speech that is entirely legitimate and even necessary in rights-respecting societies. Some kinds of expression, however, can cause real harm. It can intimidate vulnerable communities into silence, in particular when it involves advocacy of hatred that constitutes incitement to hostility, discrimination or violence. Left unchecked and viral, it can create an environment that undermines public debate and can harm even those who are not users of the subject platform. It is therefore important that States and companies address the problems of hate speech with a determination to protect those at risk of being silenced and to promote open and rigorous debate on even the most sensitive issues in the public interest.*

Given the context of online hate speech, some have suggested that it requires a multifaceted response that goes beyond civil and criminal penalties to include broader regulatory efforts and self-regulatory action by internet platforms and users.[46] It is notable that large social media providers have developed community guidelines and enforcement policies as part of their terms of use.[47]

Some countries have introduced legislation which applies directly to online hate speech. For example, Scotland's Communications Act 2003 makes it an offence to send by means of a public electronic communications network a message or other matter that is grossly offensive or of an indecent, obscene or menacing character; or cause any such message or matter to be so sent.[48] In New Zealand, the Harmful Digital Communications Act 2015 applies specifically to the online environment (although it does not apply to hate speech directed at groups of people rather than individuals) and legislation such as the Human Rights Act 1993 and the Films, Videos and Publications Classifications Act 1993 can apply to publications and electronic communications transmitted online.

However, applying (and enforcing) domestic law to online content that may be generated and held offshore has inherent difficulties. These challenges were raised by the United Kingdom Law Commission in its 2014 examination of hate crimes, including stirring up offences.[49]

---

[45] United Nations General Assembly Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression UN Doc A/74/486 (9 October 2019) at [56].

[46] This was the conclusion made by James Chalmers and Fiona Leverick "A Comparative Analysis of Hate Crime Legislation" (University of Glasgow, 2017) at 86 citing B Perry and P Olsson and "Cyberhate: the globalization of hate" (2009) 18 Information and Communications Technology Law 185 at 195-197, who note a range of legal strategies. They comment (at 196) that "law is not the only – or perhaps even the most effective – weapon available to counter cyberhate", and note "four key mechanisms": filtering, monitoring organisations, "hate speech hotlines" and internet service provider self-regulation.

[47] See, for example, YouTube's community guidelines, <https://www.youtube.com/about/policies/#community-guidelines>.

[48] Section 127(1).

[49] Law Commission, *Hate Crime: The Case for Extending the Existing Offences; History of Hate Crime Legislation* (Law Com CP No 213, Appendix B) at [2.103].

In his 2018 report to the United Nations (UN) Human Rights Council, Special Rapporteur David Kaye addressed this challenge by proposing a human rights-centred framework for the moderation of online content.[50] This included a recommendation that States introduce "smart" regulation that meets international human rights standards concerning the right to freedom of expression while focusing on transparency and remediation by companies providing online platforms.[51] He also recommended that:[52]

> *companies should recognize that the authoritative global standard for ensuring freedom of expression on their platforms is human rights law, not the varying laws of States or their own private interests, and they should re-evaluate their content standards accordingly.*

Following the events of March 15 this year, the Prime Minister has initiated the "Christchurch Call". This is an action plan that commits participating international Heads of State, Governments and leaders from the tech sector to eliminate terrorist and extremist content online. The Christchurch Call covers a range of measures including the development of tools to prevent the upload of terrorist and violent extremist content, increased transparency around the removal and detection of content and reviewing the algorithms used to detect problematic material.[53]

[50] Human Rights Council, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, UN Doc A/HRC/38/35, (6 April 2018) at [2].
[51] *Ibid*., at [66].
[52] ibid, para 70
[53] Ministry of Foreign Affairs and Trade, "Christchurch Call to Eliminate Terrorist & Violent Extremist Content Online", <https://www.christchurchcall.com/>.

Part II: International law

Under international law, speech is generally regulated by three treaties: the ICCPR, the International Convention on the Elimination of All Forms of Racial Discrimination 1965 (ICERD), and the Convention on the Prevention and Punishment of the Crime of Genocide 1948 (Genocide Convention). Each of which sets out the type of restrictions that States are either required or recommended to place on speech. These can be separated into hate speech that States *must* prohibit and hate speech that States *may* prohibit.[54] New Zealand has ratified all three of these conventions.[55]

International human rights treaties

Convention on the Prevention and Punishment of the Crime of Genocide

Direct and public incitement to genocide is prohibited in the Genocide Convention[56] and the Rome Statute of the International Criminal Court.[57] The Genocide Convention requires that States prohibit and punish as a criminal offence any "direct and public incitement to genocide," in addition to acts of genocide themselves.[58]

International Covenant on Civil and Political Rights

Article 19

Article 19(1) and (2) of the ICCPR sets out the right to freedom of expression:

1. *Everyone shall have the right to hold opinions without interference.*
2. *Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice.*

However, this is not an absolute right. Article 19(3) sets out the permissible restrictions on speech in the following terms:

3. *The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary:*

   (a) For respect of the rights or reputations of others;

   (b) For the protection of national security or of public order (ordre public), or of public health or morals.

The High Court of New Zealand in *Wall v Fairfax* has interpreted article 19 as follows:[59]

> *While the right can be restricted, the circumstances in which this is permissible are strictly limited by Article 19(3) and the restrictions must conform to the strict tests of necessity and proportionality. Specifically the restrictive measures must be appropriate to advance their protective function, they must be the least intrusive of the available measures and must be proportionate to the interest protected.*

Article 20

Article 20 of the ICCPR further limits the right to freedom of expression:

1. *Any propaganda for war shall be prohibited by law.*
2. *Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.*

---

[54] United Nations General Assembly, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, UN Doc A/66/290 (10 August 2011) at [18]. The Special Rapporteur has suggested three categories of hate speech to help identify appropriate and effective responses: (1) Hate speech that *must* be prohibited; (2) Hate speech that *may* be prohibited; (3) *Lawful* hate speech.

[55] By ratifying these treaties, New Zealand has signified an intent to be legally bound by their terms. .

[56] Article 3(c).

[57] Article 25(3)(e).

[58] Article 3(c).

[59] *Wall v Fairfax* [2018] NZHC 104 at [166.2]. This interpretation echoes the interpretation in United Nations General Assembly Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression UN Doc A/74/486 (9 October 2019).

The provision was proposed as a response to the dissemination of Nazi-Fascist propaganda.

Article 20 establishes a positive obligation on States to prohibit speech that constitutes "propaganda for war" or "advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence." States are required to "adopt the necessary legislative measures prohibiting the actions referred to therein."[60]

The three protected characteristics under article 20(2) – nationality, race, and religion – have come to be interpreted and understood as supporting the principle of equality on a larger scale. Article 2 of the ICCPR guarantees rights "without distinction of any kind" and article 26 expressly provides that "the law shall prohibit any discrimination and guarantee to all persons equal and effective protection against discrimination on any ground". The Special Rapporteur has recently noted that international standards ensure protections against adverse actions on grounds such as sex, language, religion, political opinion, sexual orientation, gender identity or intersex status, migrant or refugee status, and disability.[61] He further stated:[62]

> *Given the expansion of protection worldwide, the prohibition on incitement should be understood to apply to the broader categories now covered in international human rights law.*

The UN Human Rights Committee has made clear that article 20 does not necessarily require that hate speech be made a criminal, as opposed to a civil offence and has clarified that a prohibition under article 20 must also comply with article 19(3).[63]

Seventeen States, including New Zealand, have entered reservations to article 20 of the ICCPR. New Zealand's reservation reads:[64]

> *The Government of New Zealand having legislated in the areas of the advocacy of national and racial hatred and the exciting of hostility or ill will against any group of persons, and having regard to the right of freedom of speech, reserves the right not to introduce further legislation with regard to article 20.*

By entering reservations, States indicate that they do not intend to be bound by the article or provision to which the reservation relates. New Zealand's reservation to article 20 is significant given the absence of a "religious hatred" ground from the existing legislative framework. Sections 61 and 131 of the Human Rights Act currently only cover incitement based on colour, race, ethnic or national origins. Although article 20 requires legal prohibition of advocacy of religious hatred that constitutes incitement to discrimination, hostility or violence, the reservation that has been entered means that the New Zealand government is not obliged to legislate further to give full effect to the requirements of the article.

Article 20 also requires the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination. *Incitement of discrimination* (as opposed to incitement of hostility or contempt) is also absent from the current legislative framework in New Zealand.

---

[60] Human Rights Council, *General Comment No. 11: Prohibition of propaganda for war and inciting national, racial or religious hatred* (Art 20) U.N. Doc. CCPR/C/GC/11 (July 29 1983) at [1].

[61] United Nations General Assembly Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression UN Doc A/74/486 (9 October 2019) at [9].

[62] *Ibid*.

[63] See Human Rights Committee *Rabbae v The Netherlands*, Communication No. 2124/2011, UN Doc CCPR/C/117/D/2124/2011 (29 March 2017) at [10.4]. See also United Nations General Assembly Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression UN Doc A/74/486 (9 October 2019) at [9].

[64] See for example: Australia, Belgium, Denmark, Finland, Luxemberg, Malta, New Zealand, United Kingdom and the United States.

International Convention on the Elimination of All Forms of Racial Discrimination

Article 4 is the principal provision within ICERD for combatting racial hate speech. It requires States to make racially motivated hate speech an offence stating:

> *States Parties condemn all propaganda and all organizations which are based on ideas or theories of superiority of one race or group of persons of one colour or ethnic origin, or which attempt to justify or promote racial hatred and discrimination in any form, and undertake to adopt immediate and positive measures designed to eradicate all incitement to, or acts of, such discrimination and, to this end, with due regard to the principles embodied in the Universal Declaration of Human Rights and the rights expressly set forth in article 5 of this Convention, inter alia:*
>
> *(a) Shall declare 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, and also the provision of any assistance to racist activities, including the financing thereof;*
>
> *(b) Shall declare illegal and prohibit organizations, and also organised and all other propaganda activities, which promote and incite racial discrimination, and shall recognise participation in such organisations or activities as an offence punishable by law.*

As set out in article 4, States are to give due regard to the principles in the UDHR and the right to freedom of expression.

Unlike the ICCPR, article 4 of ICERD explicitly requires the criminalisation of speech. It applies to "all dissemination of ideas" that are racist, even if they do not involve incitement to any specified harm.

The ICERD Committee recommends that States sanction as offences punishable by law:[65]

> *(a) all dissemination of ideas based on racial or ethnic superiority or hatred, by whatever means;*
>
> *(b) incitement to hatred, contempt or discrimination against members of a group on grounds of their race, colour, descent, or national or ethnic origin;*
>
> *(c) threats or incitement to violence against persons or groups on the grounds in (b) above;*
>
> *(d) expression of insults, ridicule or slander of persons or groups or justification of hatred, contempt or discrimination on the grounds in (b) above, when it clearly amounts to incitement to hatred or discrimination; and*
>
> *(e) participation in organizations and activities which promote and incite racial discrimination.*

It also recommends the following contextual factors to be taken into account when considering whether conduct qualifies as a criminal offence: content and form of speech, economic, social and political climate, position or status of the speaker, reach of the speech, and the objective of the speech.[66]

The ICERD Committee emphasises that article 4 underlines "the international community's abhorrence of racist hate speech, understood as a form of other-directed speech[67] which rejects the core human rights principles of human dignity and equality and seeks to degrade the standing of individuals and groups in the estimation of society."[68]

[65] Committee on the Elimination of Racial Discrimination, *General Comment No 35: Combating racist hate speech*, U.N. Doc. CERD/C/GC/35 (26 September 2013) at [13].
[66] *Ibid*., at [15].
[67] "Other-directed speech" refers to speech that treats a group of people as "others" and intrinsically different from, and alien to, the speaker. It is speech based on prejudice and group identity and can contribute towards social marginalisation.
[68] *Ibid*., at [10].

New Zealand's Human Rights Review Tribunal has drawn the following general conclusions from ICERD:[69]

- Regard must be had to the principles of the UDHR, which include the right to freedom of expression.
- Criminal sanctions should be governed by the principles of legality, proportionality and necessity.
- Article 4 imposes a mandatory obligation to sanction certain offences.
- The prohibition of racist hate speech and freedom of expression should be seen as complementary and mutually supportive rights.

Three categories of speech emerge from these treaties. The first concerns hate speech that *must* be prohibited. ICCPR article 20(2) requires the restriction of any speech that constitutes advocacy of national, racial or religious hatred if it incites discrimination, hostility or violence. Article 4 of ICERD goes even further, requiring criminalisation of racist hate speech. The second refers to speech that *may* be prohibited. ICCPR article 19 allows for the restriction of speech in the interest of "respect of the rights or reputations of others" or for the protection of national security or of public order, public health or morals.

It is notable that the ICERD Committee has recommended that criminalisation of forms of racist expression should be reserved for serious cases, to be proven beyond reasonable doubt.[70] The Committee observed with concern that broad or vague restrictions on freedom of speech have been used to the detriment of groups protected by the Convention.[71]

The third category, which sits on the other side of the legal boundary, is *lawful* hate speech.[72] Lawful hate speech is expression or speech that does not give rise to criminal, civil or administrative sanctions, but still raises concern in terms of tolerance, civility and respect for the rights of others.[73] States are not required to prohibit this type of speech, however this does not mean it is morally acceptable. This point was highlighted by the judge in the most recent decision under New Zealand's hate speech laws:[74]

> *The law's limits do not define community standards or civic responsibility. I would be disappointed if anything which this Court might say could be taken as indicative of what people of one race may feel at liberty to say and which people of the other are expected to brook.*

Regional human rights instruments

The European Convention on Human Rights (ECHR), the American Convention on Human Rights, and the African Charter on Human and Peoples' Rights provide for the equivalent protection and restrictions on freedom of expression.

Article 10 of the ECHR which lists permissible restrictions on freedom of expression is longer than that under article 19(3) of the ICCPR. The European Court of Human Rights has used article 17 *Prohibition of abuse of rights* to restrict speech:

> *Nothing in this Convention may be interpreted as implying for any State, group or person any right to engage in any activity or perform any act aimed at the destruction of any of the rights and freedoms set forth herein or at their limitation to a greater extent than is provided for in the Convention.*

---

[69] *Wall v Fairfax* [2017] NZHRRT 17 at [140].

[70] Committee on the Elimination of Racial Discrimination, *General Comment No 35: Combating racist hate speech,* U.N. Doc. CERD/C/GC/35 (26 September 2013) at [12]: *The Committee recommends that the criminalisation of forms of racist expression should be reserved for serious cases, to be proven beyond reasonable doubt, while less serious cases should be addressed by means other than criminal law, taking into account, inter alia, the nature of the impact on targeted persons and groups. The application of criminal sanctions should be governed by principles of legality, proportionality and necessity.*

[71] *Ibid*., at [20].

[72] United Nations General Assembly, 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).

[73] Human Rights Council Annual report of the United Nations High Commissioner for Human Rights – Addendum 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, UN Doc. A/HRC/22/17/Add.4 (11 January 2013) Annex: Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence at [20].

[74] *Wall v Fairfax* [2018] NZHC 104 at [97] citing *Awa v Independent News* Auckland Ltd [1997] 3 NZLR 590 (CA) at 598.

The ECHR provides for restrictions where necessary for the protection of the reputation or rights of others. New Zealand courts can have regard to the ECHR, to which New Zealand is not a party, for the purpose of interpretive guidance.[75] However, regional human rights norms cannot be invoked to justify departure from international human rights protections.[76]

Other international law sources and commentary

UN human rights mechanisms and NGOs have elaborated on the meaning of treaty law on freedom of expression.

UN bodies have provided guidance on the treaty provisions. For example, in 2011 the Office of the High Commissioner for Human Rights (OHCHR) held four regional workshops on incitement to national, racial or religious hatred under article 20 of the ICCPR due to confusion about implementation. The workshops resulted in the adoption of the *Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence* (Rabat Plan). The Rabat Plan advanced a range of conclusions and recommendations for the implementation of article 20(2) and distinguished between three types of expression:[77]

- expression that constitutes a criminal offence
- expression that is not criminally punishable, but may justify a civil legal claim or administrative sanctions (for example the payment of damages)
- expression that does not give rise to criminal, civil or administrative sanctions, but still raises concern in terms of tolerance, civility and respect for the rights of others

The Rabat Plan emphasised that the threshold of the types of expression that would fall within article 20(2) should be "high and solid"[78] and that criminal sanctions should be used as a last resort. The Rabat Plan suggests a six-part threshold test for expressions considered as criminal offences: context; speaker; intent; content and form; extent of the speech act; and likelihood, including imminence.

The Rabat Plan further provides that States should ensure that the three-part test of legality, proportionality and necessity, for the restrictions to freedom of expression, apply to cases of incitement to hatred.[79]

Among the recommendations outlined in the Rabat Plan are that States should consider robust definitions of key terms such as hatred, discrimination, violence, and hostility, drawn from the guidance and definitions provided in the Camden Principles on Freedom of Expression and Equality:[80]

> *Pursuant to principle 12, national legal systems should make it clear, either explicitly or through authoritative interpretation, that the terms 'hatred' and 'hostility' refer to intense and irrational emotions of opprobrium, enmity and detestation towards the target group; the term 'advocacy' is to be understood as requiring an intention to promote hatred publicly towards the target group; and the term 'incitement' refers to statements about national, racial or religious groups which create an imminent risk of discrimination, hostility or violence against persons belonging to those groups.*

[75] *Ibid.*, at [15].
[76] United Nations General Assembly Promotion and protection of the freedom of opinion and expression UN Doc A/74/486 (9 October 2019) at [26].
[77] Human Rights Council Annual Report of the United Nations High Commissioner for Human Rights – Addendum 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, UN Doc. A/HRC/22/17/Add.4 (11 January 2013) Annex: Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence at [20].
[78] *Ibid.*, [45] and [47].
[79] *Ibid.*, [18].
[80] *Ibid.*, [21].

In addition to the terms referred to in the Rabat Plan, the organisation which developed the Camden Principles (a United Kingdom based NGO called Article 19), proposes the following key definitions:[81]

(a) *'Discrimination' shall be understood as any distinction, exclusion, restriction or preference based on race, gender, ethnicity, religion or belief, disability, age, sexual orientation, language political or other opinion, national or social origin, nationality, property, birth or other status, colour which has the purpose or effect of nullifying or impairing the recognition, enjoyment or exercise, on an equal footing, of human rights and fundamental freedoms in the political, economic, social, cultural or any other field of public life.*

(b) *'Violence' shall be understood as the intentional use of physical force or power against another person, or against a group or community that either results in or has a high likelihood of resulting in injury, death, psychological harm, maldevelopment, or deprivation.*

(c) *'Hostility' shall be understood as a manifested action of an extreme state of mind. Although the term implies a state of mind, an action is required. Hence, hostility can be defined as the manifestation of hatred – that is the manifestation of "intense and irrational emotions of opprobrium enmity and detestation towards the target group".*

The Special Rapporteur, Frank La Rue, in 2011 stated that any limitation to the right of freedom of expression must pass the following three-part, cumulative test:[82]

(a) *Legality: It must be provided by law, which is clear and accessible to everyone, with procedural safeguards, especially those guaranteed by independent courts or tribunals (principles of predictability and transparency); and*

(b) *Legitimacy: It must pursue one of the purposes set out in article 19, paragraph 3, of the Covenant, namely (i) to protect the rights or reputations of others, or (ii) to protect national security or of public order, or public health or morals (principles of legitimacy); and*

(c) *Necessity and proportionality: It must be proven as necessary and the least restrictive means required to achieve the purported aim (principles of necessity and proportionality).*

In 2019, Special Rapporteur, David Kaye, reiterated and expanded upon this three-part test.[83]

The 2011 report lists legitimate types of information which may be restricted. This includes hate speech (in order to protect the rights of affected communities), and advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence (to protect the rights of others, such as the right to life).[84] In terms of the latter, restrictions must be formulated in a way that makes clear that the sole purpose is to protect individuals and communities belonging to ethnic, national or religious groups, holding specific beliefs or opinions, whether of a religious or other nature, from hostility, discrimination or violence, rather than to protect belief systems, religions or institutions as such from criticism.[85]

---

[81] Article 19, Prohibiting incitement to discrimination, hostility or violence: December 2012 Policy Brief, (article 19, London, 2012), at 19. These definitions were adapted from the definition of discrimination advanced by the *Convention on the Elimination of All Forms of Discrimination against Women and the Convention on the Elimination of All Forms of Racial Discrimination* and the definition of violence by the World Health Organisation in the 2002 report *World Report on Violence and Health*.

[82] Human Rights Council, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Frank La Rue*, U.N. Doc A/HRC/17/27 (16 May 2011) at [24].

[83] United Nations General Assembly Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression UN Doc A/74/486 (9 October 2019) at [6].

[84] Ibid., at [25]. The issue of hate speech has also been addressed in previous reports, see inter alia, UN Documents: E/CN.4/1999/64; E/CN.4/2000/63; E/CN.4/2002/75; and A/HRC/4/27.

[85] Human Rights Council, *Annual report of the United Nations High Commissioner for Human rights: 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*, U.N. Doc A/HRC/22/17/Add.4 (Jan. 11, 2013) at [11].

The Special Rapporteurs have provided further clarity on speech which is prohibited under article 20(2).[86] First is that only advocacy of hatred is covered. Second is that it must constitute incitement to one of the three listed results. Third is that the incitement is likely to result in discrimination, hostility or violence. Therefore, advocacy of hatred only becomes a breach when the speaker seeks to provoke reactions on the part of the audience and there is a very close link between the expression and the resulting risk of discrimination, hostility or violence.[87]

The Human Rights Council made clear in Resolution 12/16 that certain types of speech should never be subject to restrictions. These categories include discussion of government policies and political debate; reporting on human rights, government activities and corruption of government; engaging in election campaigns, peaceful demonstrations or political activities, including for peace or democracy; and expression of opinion and dissent, religion or belief, including by persons belonging to minorities or vulnerable groups.[88]

In 2019 the Special Rapporteur provided further examples of speech that should not be restricted. First, where ridicule, abuse, criticism or other 'attacks' is seen as offensive to religious or other belief systems but does not constitute incitement. Second, opinions that are "erroneous" or "incorrect interpretations of past events". Third, where a speaker is individually targeting an identifiable victim but not seeking to incite others to take an action against persons based on protected characteristics. Fourth, expression that may be offensive or characterised by prejudice and raises serious concerns of intolerance but does not meet a threshold of severity to merit any kind of restriction. In the latter, the Special Rapporteur recommends that the six factors identified by the Rabat Plan provide a rubric for considering how to evaluate the appropriate response to such speech.

The Special Rapporteur emphasised that the absence of restriction under international law does not mean an absence of action. States should take robust steps to counter such intolerance and ensure that public authorities protect individuals against discrimination rooted in these kinds of assertions of hate. Examples include Government condemnation of prejudice, education, training, public service announcements, and community projects.[89]

In September 2019 a joint open letter signed by 26 United Nations experts raised concerns about the global increase in hate speech.[90] Signatories included the UN Special Rapporteur on the Right to Freedom of Expression and Opinion, the Special Rapporteur on Freedom of Religion and Belief, the Independent Expert on Protection Against Violence and Discrimination Based on Sexual Orientation and Gender Identify and the Special Rapporteur on Contemporary Forms of Racism, Racial Discrimination, Xenophobia and Related Intolerance. The letter states:

[86] United Nations General Assembly, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, U.N. Doc A/66/290 (Aug 10, 2011); United Nations General Assembly Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression UN Doc A/74/486 (9 October 2019) at [8].

[87] United Nations General Assembly, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, U.N. Doc A/66/290 (Aug 10, 2011) at [28].

[88] Human Rights Council, *Resolution 12/16 Freedom of option and expression*, UN Doc A/HRC/RES/12/16 (12 October 2009) at [5(p)]; United Nations General Assembly, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, U.N. Doc A/66/290 (Aug 10, 2011) at [42].

[89] United Nations General Assembly Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression UN Doc A/74/486 (9 October 2019) at [21]-[24].

[90] OHCHR, Joint open letter on concerns about the global increase in hate speech <www.ohchr.org>.

> *We are alarmed by the recent increase in hateful messages and incitement to discrimination and hatred against migrants, minority groups and various ethnic groups, as well as the defenders of their rights, in numerous countries. Hate speech, both online and offline, has exacerbated societal and racial tensions, inciting attacks with deadly consequences around the world.*

The letter calls for the rhetoric of hatred to be countered as it has real-life consequences, referring to the correlation between hate speech and the number of hate crimes committed. The experts also warn that the demonisation of entire groups of people as dangerous or inferior is not new in history and has led to catastrophic tragedies in the past.

Significantly, the signatories reiterate the importance of freedom of expression and its role as a "vital tool" to counter hate speech. They encourage the application of the three-part test for restrictions to freedom of expression - legality, proportionality and necessity - to cases of incitement to hatred and also urge States to follow the standards adopted in the Rabat Plan.

The letter concludes with a call for States to:

> *……actively work towards policies that guarantee the rights to equality and non-discrimination and freedom of expression, as well as the right to live a life free of violence through the promotion of tolerance, diversity and pluralistic views; these are the centre of pluralistic and democratic societies. We believe that these efforts will help make countries safer, and foster the inclusive and peaceful societies that we would all like and deserve to live in.*

Part III: Hate speech laws in New Zealand

Human Rights Act

The racial disharmony provisions contained in sections 61 and 131 of the Human Rights Act 1993 (HRA) are generally referred to as New Zealand's hate speech laws. The predecessors to these provisions were enacted to meet New Zealand's obligations under the ICERD set out in the previous section.[91] The HRA provides both civil and criminal remedies for hate speech, but only in relation to hate speech directed at race, colour, ethnic or national origins. Section 61 provides for a civil law remedy for racial disharmony and section 131 provides for a separate criminal offence of inciting racial disharmony (see Appendix 1 for the laws in full).

Section 61: Racial disharmony

Racial disharmony is a form of discrimination in the public sphere directed against a group of persons, rather than an individual. Section 61 of the HRA makes it unlawful for any person to publish, broadcast or distribute written matter or use words in public which are "threatening, abusive, or insulting" and likely to "excite hostility against or bring into contempt any group of persons…on the ground of the colour, race, or ethnic or national origins of that group of persons." Section 61 also covers material that has been broadcast by electronic communication which means the provisions can apply to the on-line environment.

The New Zealand Human Rights Commission (Commission) can receive complaints about alleged breaches of section 61. The Commission can provide dispute resolution assistance to help the parties to the complaint resolve the concerns that have been identified. This includes providing a free, independent mediation service to members of the public who have made complaints.

If a complaint is not resolved through the Commission's processes, the complainant can take the complaint to the Human Rights Review Tribunal, an independent judicial body. The Human Rights Review Tribunal has jurisdiction to provide remedies including damages, declarations and the issuing of training orders if it determines that section 61 has been breached. Complainants can also approach the Office of Human Rights Proceedings, a separate and independent division of the Commission, to request free legal assistance with pursuing claims before the Human Rights Review Tribunal. In the 2017/2018 year the Director of the Office of Human Rights Proceedings provided assistance to 31 percent of the applicants who sought assistance.

Between 2014 and 2018, the Commission received 218 complaints of racial disharmony. These do not represent individual instances giving rise to a complaint. For example, 37 of the complaints received in 2018 related to one newspaper article. Only one complaint under section 61 has been considered by the Human Rights Review Tribunal.

The original section 61 first appeared in 1977 as section 9A of the Race Relations Act 1971, which was an Act intended to affirm and promote racial equality in New Zealand and to implement ICERD. The then Minister of Justice, Hon David Thomson, described the proposed section 9A as importing "the milder processes of conciliation and the civil law to deal with cases where the language used was not sufficiently flagrant to lend itself readily to criminal prosecution."[92] It was intended that this section would provide additional power for the then Race Relations Conciliator to combat racial prejudice.[93]

[91] *Wall v Fairfax* [2017] NZHRRT 17; *Wall v Fairfax* [2018] NZHC 104.
[92] *Wall v Fairfax* [2017] NZHRRT 17 at [119].
[93] *Ibid.*, at [119].

In 1989 section 9A of the Race Relations Act was repealed as it was not working as intended. An explanation for the repeal is provided in *Brookers Human Rights Law*: the wording allowed the media to be prosecuted for reporting material leading to the exciting of racial disharmony but exempted those who made the comments if they did so in a private place.[94] However, in 1993 when the new HRA was enacted, section 9A was revived as section 61, with some changes. The reference to exciting "ill-will" or bringing people into "ridicule" was removed and the media exemption in section 61(2) was added (see Appendix 1).

In 1996, the then Complaints Review Tribunal (later renamed the Human Rights Review Tribunal) considered the implications of section 61 in the case of *Proceedings Commissioner v Archer*. The Tribunal held that words used in a radio broadcast were likely to excite hostility or bring into contempt Chinese and Japanese people living in New Zealand on the grounds of their colour, race or national or ethnic origins.[95] A reasonable person test was found to be the appropriate measure in determining whether the words were threatening, abusive or insulting.

In 2004 the Government Administration Select Committee initiated an inquiry into hate speech.[96] The inquiry considered how the legislation had worked in practice and whether changes were needed, including whether the law should be extended to cover inciting hatred against people on the grounds of their religion, gender or sexual orientation.[97] The terms of reference for the inquiry included consideration of:[98]

- whether or not further legislation to prohibit or restrain hate speech was warranted;
- whether censorship of material that vilifies certain groups would be a justified limitation on the rights and freedoms affirmed by the New Zealand Bill of Rights Act 1990;
- an appropriate threshold test for prohibition or restraint of hate speech;
- whether any prohibition or restraint of hate speech or hateful expressions would be a justified limitation on the rights and freedoms outlined in the New Zealand Bill of Rights Act 1990; and
- the steps taken by the international community to control hate speech and hateful expressions.

In late 2004, the Committee received submissions on the matter. The Office of Film and Literature Classification, the Independent Crown Entity responsible for the classification of "publications", submitted that the New Zealand legislation in place at that time did not specifically or effectively address "hate speech".[99] It further submitted that there was substantial precedent for a limited restriction on the freedom of expression to remedy the social harm caused by speech that incites hatred against individuals and groups on the basis of characteristics that are already prohibited grounds of discrimination, including race, ethnicity, colour, nationality, religion and sexual orientation.[100] Unfortunately, there does not appear to be a report of the Committee on the Inquiry into Hate Speech but some of the matters raised by the Office of Film and Literature Classification were partially addressed

[94] *Ibid*., at [119] citing Bell (ed) *Brookers Human Rights Law* (loose leaf ed, Thomas Reuters) at HR61.01.
[95] *Proceedings Commissioner v Archer* (1996) 3 HRNZ 123.
[96] Beehive, Goff welcomes hate speech inquiry (7 August 2004) <www.beehive.govt.nz>; see also Government Administration Committee, Films, Videos and Publications Classification Amendment: Commentary (House of Representatives, Wellington 2004) at 3 which stated: *We considered carefully whether to widen the meaning of ''objectionable'' in section 3 of the Act to include hate speech and concluded it was beyond the policy of this bill. The bill primarily caters, in terms of classification, for the proliferation of child sex abuse images via the Internet. Hate speech raises wider legal issues, including the fundamental right in a democracy to freedom of expression. In New Zealand this freedom may be subject to reasonable limits under section 5 of the New Zealand Bill of Rights Act 1990. Section 3 is a specific example of such limits. We were mindful of the need to be cautious in placing further limitations on freedom of expression, however well-meaning, without very careful scrutiny to ensure that any limitation is reasonable and not open to exploitation. Hate speech also falls within the right to freedom from discrimination, and will require further consideration of human rights law. Therefore, the committee has initiated, under Standing Order 189(2), an inquiry into hate speech. Separately, the Minister of Justice has advised us that he will refer the topic to the Law Commission for further study. We anticipate this study and our inquiry will complement each other to provide a sound basis for the determination of these difficult issues.*
[97] Beehive, Goff welcomes hate speech inquiry (7 August 2004)) <www.beehive.govt.nz>.
[98] Human Rights Commission, *Race Relations* in 2004, (March 2005) www.hrc.co.nz at [6.3].
[98] Office of Film & Literature Classification, Submission: Inquiry into Hate Speech (29 October 2004) <www.classificationoffice.govt.nz > at [3] and [27].
[100] *Ibid*.

subsequently in the Harmful Digital Communications Act, which is discussed further below.

The scope and application of section 61 was recently considered by the High Court in *Wall v Fairfax*.[101] The Court concluded that cartoons published by the defendants, although offensive, were not likely to excite hostility or contempt at the level of abhorrence, delegitimisation and rejection that could realistically threaten racial disharmony in New Zealand.[102] The Court accordingly held that the publication was not captured by the section.[103] In coming to its conclusion, the Court made some useful observations regarding section 61:

- The civil remedy for racial disharmony is directed at the prevention of discrimination in the form of racist speech, the promotion of racial harmony and to meet New Zealand's obligations under ICERD.[104]
- Section 61 establishes a high threshold, targeted to racist speech at the serious end of the spectrum that "applies only to relatively egregious examples of expression which inspire enmity, extreme ill-will or are likely to result in the group being despised."[105]
- Section 61 applies only to material that relates to race, colour, ethnic or national origins. Material aimed at individuals or groups because of their religion is out of scope.
- A two-part test must be satisfied. First, the expression must be "threatening, abusive, or insulting." Second, it must be "likely to excite hostility or bring into contempt any group of persons" in New Zealand on the ground of the colour, race, or ethnic or national origins of that group.
- "Excite hostility" or "bring into contempt" involves an objective test – "whether a reasonable person, aware of the context and circumstances surrounding the expression, would view it as likely to expose the protected group to the identified consequence."[106]
- The focus is on the effect of the words on others outside of the group, with reference to context and circumstances.[107] The Court considered who must be likely to be excited to hostility or contempt and found that the focus should be on those who are "susceptible" or "persuadable". The Court interpreted the verbs "excite" and "bring" to connote a change in behaviour or thinking requiring consideration of whether such people are likely to become hostile or contemptuous as a result of the words used.[108]

Section 131: Inciting racial disharmony

Section 131 provides a criminal sanction for inciting racial disharmony. Section 131 largely repeats the same test as for section 61, except that an accused must have "intended" that his or her conduct would have the effect of inciting hostility or ill-will towards a specified group or bringing that group into contempt or ridicule. The crime attracts potential penalties of a term of imprisonment not exceeding three months or a fine not exceeding $7,000. Section 132 of the HRA requires the consent of the Attorney General to be obtained before a prosecution can be instituted under section 131.

Section 131 was originally included in legislation as section 25 of the Race Relations Act 1971.[109] The only prosecution under section 25 or its successors, is the 1979 Court of Appeal case of *King-Ansell v Police*.[110] In

[101] *Wall v Fairfax* [2018] NZHC 104.
[102] *Ibid*., at [94].
[103] *Ibid*., at [94].
[104] See parliamentary comment on the introduction of s 9A into the Race Relations Act (20 July 1977) 411 NZPD 1477.
[105] *Wall v Fairfax* [2018] NZHC 104 at [42] and [52]. The case law is clear that the threshold that must be reached to establish a breach is extremely high. Referring to New Zealand's obligations under article 4 of the Convention on the Elimination of Racial Discrimination ICERD, the Court held: *In our view, the Tribunal was correct to conclude that both civil and criminal conduct caught by art 4 will be "at the serious end of the spectrum". That inference is warranted by the relatively strong language of art 4 ("ideas or theories of superiority of one race or group of persons"… "racial hatred and discrimination"), which we do not read as referring to low level insulting speech, It is further warranted by the explicit reference to the principles embodied in the UDHR, which requires due regard to be had to freedom of speech when implementing art 4.*
[106] *Ibid*., [51].
[107] *Ibid*., [50].
[108] *Ibid*., [63].
[109] *Wall v Fairfax* [2017] NZHRRT 17 at [118].
[110] *King- Ansell v Police* [1979] 2 NZLR 531.

that case, King-Ansell was charged with, and convicted of, publishing a pamphlet which incited ill-will against Jewish people on the grounds of their ethnic origins. On appeal, the Court held that "ethnic" should be taken to mean "pertaining to race or nation." In the decision, Woodhouse J and Richardson J stated that section 25 referred to discrimination against a group, not an individual:

> *The ultimate genetic ancestry of any New Zealander is not susceptible to legal proof. Race is clearly used in its popular meaning. So are the other words. The real test is whether the individuals or the group regard themselves and are regarded by others in the community as having a particular historical identity in terms of their colour or their racial, national or ethnic origins. That must be based on a belief shared by members of the* **group.**[111]

Harmful Digital Communications Act

Online hate speech in New Zealand directed at individuals is principally regulated through the Harmful Digital Communications Act (HDCA). The HDCA sets out ten communication principles. Principle 10 reads:

> *A digital communication should not denigrate an individual by reason of his or her colour, race, ethnic or national origins, religion, gender, sexual orientation, or disability.*

Both the scope and threshold of the HDCA are different to those included in sections 61 and 131 of the HRA. The protected characteristics in the HDCA include religion, gender, sexual orientation and disability in addition to the HRA's colour, race, ethnic or national origins. Neither is it necessary under the HDCA to show incitement of third parties. Only denigration of an individual on a specified ground needs to be established.

Complaints can be made to Netsafe, the approved agency under the HDCA, if an individual believes that one of the communication principles has been breached. Netsafe will work with the parties to find a resolution. However, if a resolution cannot be reached the agency must notify the complainant of their right to apply to the District Court for an order if the affected individual believes he or she has suffered, or will suffer, harm as a result of the digital communication concerned. The Court can make a range of orders including requiring the removal of content and the publishing of an apology. The HDCA also provides for criminal liability when a person does not comply with an order or when a person posts a digital communication with the intention that it cause harm and then harm results.

In the recent District Court decision of *R v Iyer* it was held that the term "posts a digital communication" is broad.[112] Digital communications are not confined to only "one on one" communications but can include everything in the realm of cyberspace that has the capability of being published and viewed.[113] This could be seen to include content on social media platforms such as YouTube, Facebook, and Twitter.

The HDCA requires online content hosts to take certain steps after receiving a notice of complaint. The steps require the on-line content host to notify the user who posted the harmful content and to take down the content unless a valid counter-notice is received by the user within 48 hours. If an online content host takes these steps, they are protected against civil or criminal liability for hosting the harmful content.

Films, Videos and Publications Classification Act

The Films, Videos and Publications Classification Act 1993 (FVPC Act) allows a complaint to be made where a publication is deemed "objectionable" and thus is also relevant to the discussion on hate speech. The definition of a "publication" under the FVPC Act is broad and can apply to images, representations, signs, statements, or words that are online. However, the FVPC Act currently exempts from liability network operators or online service providers who provide the network or facility through which objectionable publications are distributed.

[111] *Ibid.*, Woodhouse J at 537 and Richardson J at 542 (emphasis added).
[112] *R v Partha Iyer* [2016] NZDC 23957 at [29].
[113] *Ibid.*, at [37].

The definition of "objectionable" at section 3(3) includes where the publication represents (whether directly or by implication) that members of any particular class of the public are inherently inferior to other members of the public by reason of any characteristic of members of that class, because of a characteristic that is a prohibited ground of discrimination under section 21(1) of the HRA. This section does not align directly with section 61 of the HRA and is wider in scope, covering representations of inferiority and not requiring any incitement element.

Broadcasting Act

The Broadcasting Act 1989 requires broadcasters to maintain standards that are consistent with good taste and decency and to ensure that controversial issues of public importance are discussed in a balanced way. Broadcasting standards have been developed in eleven areas, including discrimination and denigration. The relevant standard reflects section 21(1)(c) (iv) of the legislation which requires safeguards against the portrayal of persons in a manner that encourages denigration of, or discrimination against, sections of the community on account of sex, race, age, disability or occupation status or as a consequence of legitimate expression of religious, cultural or political beliefs.

Summary Offences Act

The Summary Offences Act 1981 includes sections on inciting or encouraging disorderly behaviour and provisions relating to offensive behaviour and language. Section 3 states that people who, in or within view of any public place, behave, or incite or encourage any person to behave in a riotous, offensive, threatening, insulting or disorderly manner that is likely in the circumstances to cause violence against persons or property to start or continue, is liable to imprisonment or a fine.[114] Section 4 states that a person is liable to a fine where in any public place or within hearing of a public place, that person uses words to alarm, insult or offend or addresses any indecent or obscene words to any person.[115] Depending on the circumstances, these provisions can have application in situations involving hate speech.

UN recommendations

There have been several recommendations by UN human rights monitoring bodies relating to hate speech in New Zealand, including recommendations made during the Universal Periodic Review process and following the review of New Zealand by the Committee on the Elimination of Racial Discrimination.

The Universal Periodic Review of New Zealand in 2019 included several recommendations relating to combatting racism, discrimination, xenophobia and hate crime.[116] These recommendations were all accepted by the New Zealand government.[117] They included recommendations that the government:

> *Continue efforts to combat racial discrimination and hate speech and promote diversity and tolerance* (Recommendation 44); *and*
>
> *Develop and implement a national plan of action against racial discrimination, xenophobia and hate crime* (Recommendation 48)

[109] *Wall v Fairfax* [2017] NZHRRT 17 at [118].
[110] *King- Ansell v Police* [1979] 2 NZLR 531.
[111] *Ibid.*, Woodhouse J at 537 and Richardson J at 542 (emphasis added).
[112] *R v Partha Iyer* [2016] NZDC 23957 at [29].
[113] *Ibid.*, at [37].
[114] Imprisonment for a term not exceeding 3 months or a fine not exceeding $2,000.
[115] Section 4(1) states: *Every person is liable to a fine not exceeding $1,000 who,—(a) in or within view of any public place, behaves in an offensive or disorderly manner; or (b) in any public place, addresses any words to any person intending to threaten, alarm, insult, or offend that person; or (c) in or within hearing of a public place,—(i) uses any threatening or insulting words and is reckless whether any person is alarmed or insulted by those words; or (ii) addresses any indecent or obscene words to any person. Section 4(3) clarifies that in determining whether words were indecent or obscene, the court shall have regard to all the circumstances pertaining at the material time, including whether defendant had reasonable grounds for believing that the person to whom the words were addressed, or any person by whom they might be overheard, would not be offended. Aspects of the Crimes Act 1961, including incitement to commit an offence under section 66 are also relevant where hate speech incites the commitment of an offence.*
[116] Human Rights Council, Report of the Working Group on the Universal Periodic Review: New Zealand, UN Doc A/HRC/41/4 (1 April 2019) at [122.44]-[122.48].
[117] Human Rights Council, *Report of the Working Group on the Universal Periodic Review: New Zealand – Addendum Views on conclusions and/or recommendations, voluntary commitments and replies presented by the State under review*, UN Doc A/HRC/41/4/Add.1 (17 June 2019) at [17]: (Recommendations 44-48).

The 2014 Universal Periodic Review also addressed hate speech, recommending that the New Zealand government "continue combatting and condemning racism and hate speech used by politicians and the expression of racism in the media, in particular discriminatory language and hate speech and the dissemination of racist ideas and languages".[118] This recommendation was accepted by the New Zealand government.[119]

The Committee on the Elimination of Racial Discrimination has recommended review of, or changes to, New Zealand's hate speech provisions on a number of occasions. For example, in its 2017 report on New Zealand, the Committee made recommendations relating to hate speech.[120] Recommendation 2 stated:

> *Recalling its general recommendation No. 35 (2013) on combating racist hate speech, the Committee recommends that the State party:*
>
> *(a) Review the adequacy of current legislation in addressing and sanctioning racist hate speech and incitement to racial hatred, and ensure that the legislative framework conforms to article 4 of the Convention.*

Recommendation 3 further stated:

> *Recalling its general recommendation No. 35 (2013) on combating racist hate speech, the Committee recommends that the State party:*
>
> *Ensure that all incidents of racist hate crimes and racist hate speech are investigated and prosecuted, that the perpetrators are sanctioned and that victims are compensated.*

The Committee has also recommended that New Zealand make the optional declaration provided for in article 14 of the ICERD, recognising the Committee's competence to receive and consider communications from individuals or groups claiming to be victims of violations of any of the rights in the Convention.[121] New Zealand has not yet made a declaration under article 14.[122]

Overview of hate speech laws in New Zealand

The hate speech laws in New Zealand are limited in scope.

The HRA provides both civil and criminal remedies for hate speech, but only in relation to hate speech directed at race, colour, ethnic or national origins. The full range of protected characteristics listed in the HRA are not covered under the current hate speech provisions in the HRA. Characteristics such as disability, sex, sexual orientation, religious belief, age, family status, employment status and political opinion are absent.

The HDCA makes it illegal to post certain types of digital communications that cause harm across a broader range of protected characteristics and has a different legal test. However, this legislation only applies to online hate speech. While these provisions initially only applied to one-on-one communications and not to hate speech against groups, recent case law has interpreted the legislation more broadly.

It is also noted that if a criminal offence is found to be motivated by hate against an individual on the basis of a protected characteristic, then this is regarded as an aggravating factor in sentencing. However, the existence of a hate-based motivation factor for sentencing purposes is not equivalent to a stand-alone criminal offence.

While there are some provisions of the FVPC Act and the Broadcasting Act that are relevant to the discussion about hate speech, these are limited to the scope of their respective legislation, namely as it relates to film, videos and publications and broadcasting.

Similarly, the Summary Offences Act may be relevant depending on the circumstances.

---

[118] Human Rights Council, Report of the Working Group on the Universal Periodic Review: New Zealand, UN Doc A/HRC/26/3 (7 April 2014) at [122.100].

[119] Human Rights Council, *Report of the Working Group on the Universal Periodic Review: New Zealand – Addendum Views on conclusions and/or recommendations, voluntary commitments and replies presented by the State under review*, UN Doc A/HRC/26/3/Add.1 (26 May 2014) at [15].

[120] Committee on the Elimination of Racial Discrimination, Concluding observations on the combined twenty-first and twenty-second period reports of New Zealand, UN Doc. CERD/C/NZL/CO/21-22 (22 September 2017) at [9] and [11].

[121] At [42].

[122] Ministry of Justice, "Constitutional Issues and Human Rights" <https://www.justice.govt.nz/justice-sector-policy/constitutional-issues-and-human-rights/human-rights/international-human-rights/cerd/>; United Nations Treaty Collection, "Status of Treaties: International Convention on the Elimination of All Forms of Discrimination" <https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-2&chapter=4&clang=_en>.

Part IV: Hate speech laws in other countries

National hate speech laws vary in countries that New Zealand often compares itself to. In some countries, incitement to hatred is a criminal offence, while in others, it is an offence under both criminal and civil law or under civil law only. This part of the paper will outline national hate speech laws in Australia, Canada, the United Kingdom, Ireland and the United States.

Annex 2 of this paper includes a table comparing what characteristics are protected under national criminal hate speech laws in the various countries.

Australia

Australia's laws that prohibit hate speech are known as "vilification" or "serious vilification" laws. Anti-vilification laws exist federally, in all six states and the Australian Capital Territory (ACT). The focus in Australia is mainly on civil, rather than criminal, anti-vilification laws.

Under the civil anti-vilification laws, if a person feels that an incident of unlawful vilification has occurred, they can lodge a complaint to the relevant state authority, such as the New South Wales Anti-Discrimination Board or the Australian Human Rights Commission (AHRC). After a person lodges a complaint, the authority will assess whether the allegation falls within the definition of unlawful vilification. If so, the authority will investigate and attempt to mediate a resolution. Around 200 complaints are lodged each year and less than two percent of those complaints end up in a court or tribunal.

Federal

Australia does not have a federal law criminalising hate speech. However, a civil remedy is provided under the Racial Hatred Act 1995 based on race, colour, or national or ethnic origin. Section 18C(1) of the Act provides:

> *It is unlawful for a person to do an act, otherwise than in private, if: (a) the act is reasonably likely, in all the circumstances, to offend, insult, humiliate or intimidate another person or a group of people; and (b) the act is done because of the race, colour or national or ethnic origin of the other person or of some or all of the people in the group.*

A person can make a complaint to the AHRC about unlawful acts under section 18C(1)[124] and the AHRC will attempt to conciliate the matter. If the AHRC cannot negotiate an agreement which is acceptable to the complainant, the complainant's only redress is through the Federal Court or through the Federal Magistrates Service.

In 2017, the Turnbull government introduced legislation in an effort to replace the words "offend, insult or humiliate" in section 18C(1) with the term "harass". A majority of the senate rejected the legislative proposal and the section remains in its original form today.

State and territorial

This section sets out the law in each of Australia's six states and two territories. Four of the states and territories have civil and criminal hate speech laws (New South Wales, Queensland, Victoria, and the ACT). The protected characteristics vary between states and territories.

New South Wales

New South Wales law provides for both civil and criminal anti-vilification laws. The civil remedy is set out in the Anti-Discrimination Act 1977 (ADA) which makes

[123] Katharine Gelber and Luke McNamara "Anti-vilification Laws and public racism in Australia: Mapping the gaps between the harms occasioned and the remedies provided" (2016) 39 *UNSW Law Journal* 488 at 494; Gelber and McNamara, Katharine Gelber and Luke McNamara, 'Private Litigation to Address a Public Wrong: A Study of Australia's Regulatory Response to "Hate Speech"' (2014) 33 Civil Justice Quarterly 307 at 314.

[124] Australian Human Rights Commission Act 1986.

racial vilification,[125] transgender vilification[126] and HIV/AIDs vilification[127] unlawful. Under each of these sections it is unlawful to "incite hatred towards, serious contempt for, or severe ridicule" of a person or group of persons on the grounds of race, transgender, or HIV/AIDS status. There has not been a prosecution under this section for 30 years.[128]

Up until 2018 the ADA also provided for the criminal offences of serious racial vilification,[129] serious transgender vilification,[130] serious homosexual[131] vilification[132] and serious HIV/AIDs vilification. However, these crimes were recently moved to the Crimes Act 1900 following the passing of the Crimes Amendment (Publicly Threatening and Inciting Violence) Bill 2018 in June 2018.[133] In addition, the grounds were broadened to include religion. Section 93Z of the Crimes Act now sets out the criminal offence of publicly threatening or inciting violence on the grounds of race, religion, sexual orientation, gender identity or intersex or HIV/AIDS status to create a new criminal offence where:

> *A person who, by a public act, intentionally or recklessly threatens or incites violence towards another person or a group of persons on any of the following grounds is guilty of an offence…*

The penalty was also increased to a maximum term of three years imprisonment.

The removal of criminal offences from the ADA was a result of an inquiry on racial vilification law in New South Wales which was established by the New South Wales Parliament in 2012. The Standing Committee on Law and Justice produced its report a year later.[134]

In June 2018, the New South Wales government provided its response to the report of the inquiry and indicated that it would introduce a Bill to Parliament to implement the following proposed reforms:[135]

- Broadening grounds to include 'religious belief or affiliation, or absence thereof'
- Dealing with 'threatening violence 'in addition to 'inciting violence'
- Moving all serious vilification offences from the ADA into the Crimes Act 1900

Victoria

Victoria has both civil and criminal vilification laws under the Racial and Religious Tolerance Act 2001 on the grounds of race and religion.

Under the Act it is unlawful to engage in conduct that "incites hatred against, serious contempt for, or revulsion or severe ridicule of, that other person or class of persons" on the grounds of race or religious belief or activity.[136] A person can lodge a complaint with the Victoria Human Rights Commission and a decision will be made whether to conciliate the complaint. The complainant can make a complaint to the Victorian Civil and Administrative Tribunal if the complaint cannot be resolved.[137]

The Act also provides for separate criminal offences of serious racial vilification and serious religious vilification. A person must not, on the ground of race or religious belief or activity:[138]

[125] Anti-Discrimination Act 1977 (NSW) s 20(c).
[126] Anti-Discrimination Act 1977 (NSW) s 38S.
[127] Anti-Discrimination Act 1977 (NSW) s 49ZXB.
[128] Stepan Kerkyasharian Ao, Report on Consultation: Serious Vilification Laws in NSW (May 2017), at 6; See also Christopher Knaus, "Tougher race hate laws for NSW as government reverses position again" (5 June 2018) The Guardian <www.theguardian.com>.
[129] Anti-Discrimination Act 1977 (NSW) s 20D.
[130] Anti-Discrimination Act 1977 (NSW) s 38T, as inserted by the Transgender (Anti-Discrimination and Other Acts Amendment) Act 1996 (NSW) s 3.
[131] Anti-Discrimination Act 1977 (NSW) s 49ZTA, as inserted by the Anti-Discrimination (Homosexual Vilification) Amendment Act 1993 s 3.
[132] Anti-Discrimination Act 1977 (NSW) s 49ZXC, as inserted by the Anti-Discrimination (Amendment) Act 1994 s 3.
[133] Crimes Amendment (Publicly Threatening and Inciting Violence) Bill 2018 [NSW], Passed by both Houses (2018).
[134] Legislative Council, Standing Committee on Law and Justice, *Racial vilification law in New South Wales* – Report 50 (3 December 2013).
[135] Mark Speakman, Attorney General, Government Response to the Legislative Council Standing Committee on Law and Justice's Inquiry into Racial Vilification law in NSW, (5 June 2018).
[136] Racial and Religious Tolerance Act 2001, ss 7 and 8.
[137] Equal Opportunity Act 1995.
[138] Racial and Religious Tolerance Act 2001, ss 24-25.

> *. . . intentionally engage in conduct that the offender knows is likely – (a) to incite hatred against that other person or class of persons; and (b) to threaten, or incite others to threaten, physical harm towards that other person or class of persons or the property of that other person or class of persons.*

The penalty for an individual is imprisonment for six months and/or 60 penalty units and in the case of a body corporate 300 penalty units.[139]

Queensland

Like Victoria and New South Wales, Queensland has both civil and criminal vilification laws under the Anti-Discrimination Act 1991.

A civil remedy for vilification on the grounds of race, religion, sexuality or gender identity is set out under section 124A of the Act:

> *A person must not, by a public act, incite hatred towards, serious contempt for, or severe ridicule of, a person or group of persons on the ground of the race, religion, sexuality or gender identity of the person or members of the group.*

An individual can make a complaint to the Queensland Anti-Discrimination Commission. If the complaint is accepted, the Commission will investigate and attempt to resolve through conciliation. Where complaints cannot be conciliated the complainant can seek referral to a tribunal.[140]

Section 131A of the Act also provides for the criminal offences of serious vilification on the grounds of race, religion, sexuality or gender:

> *(1) A person must not, by a public act, knowingly or recklessly incite hatred towards, serious contempt for, or severe ridicule of, a person or group of persons on the ground of the race, religion, sexuality or gender identity of the person or members of the group in a way that includes— (a) threatening physical harm towards, or towards any property of, the person or group of persons; or (b) inciting others to threaten physical harm towards, or towards any property of, the person or group of persons.*

The offence carries a penalty of up to six months imprisonment and or 70 penalty units. Written consent of a Crown Law Officer is required before a proceeding is commenced under this section.[141]

South Australia

South Australia has no civil vilification law. However, the Racial Vilification Act 1996 makes racial vilification a criminal offence.[142] Section 4 provides:

> *A person must not, by a public act, incite hatred towards, serious contempt for, or severe ridicule of, a person or group of persons on the ground of their race by—*
>
> *(a) threatening physical harm to the person, or members of the group, or to property of the person or members of the group; or*
>
> *(b) inciting others to threaten physical harm to the person, or members of the group, or to property of the person or members of the group.*

The offence carries a penalty of imprisonment for three years and/or a $5,000 fine for an individual. A prosecution for the offence cannot be commenced without the Director of Public Prosecutions' written consent.[143]

[139] At ss 24-25.
[140] Work-related complaints are dealt with by the Queensland Industrial Relations Commission. All other complaints are dealt with by the Queensland Civil and Administrative Tribunal.
[141] Anti-Discrimination Act 1991, s 131A(2).
[142] Racial Vilification Act 1996.
[143] Racial Vilification Act 1996, s 5

Western Australia

Western Australia has no civil vilification law. However, the Criminal Code 1913 sets out the crimes of racial harassment and incitement to racial hatred. There are a range of specific offences relating to incitement to racial hatred.[144] The offences carry maximum imprisonment terms between five and 14 years. Or on summary conviction a penalty of two years and a maximum fine of $24,000.

Tasmania

Tasmania does not have any vilification laws. However, the Tasmania Anti-Discrimination Act 1998 lists "inciting hatred" as a prohibited conduct. Section 19 provides that a person, by a public act, must not incite hatred towards, serious contempt for, or severe ridicule of, a person or a group of persons on the ground of race, disability, sexual orientation or lawful sexual activity, religious belief or affiliation or activity or the gender identity or intersex variations of sex characteristics.[145]

Anyone can make a complaint to the Anti-Discrimination Commissioner who will attempt to resolve by conciliation or in any other way.[146]

Australian Capital Territories

The ACT has both civil and criminal vilification laws. A civil remedy is provided under section 67A of the Discrimination Act 1991 which makes it unlawful "for a person to incite hatred toward, revulsion of, serious contempt for, or even severe ridicule of a person or group of people on the ground of" disability, gender identity, HIV/AIDS status, intersex status, race, religious conviction, or sexuality.[147]

The Criminal Code 2002 provides for the criminal offence of serious vilification:

> *A person commits an offence if they intentionally carry out a threatening act, the act is threatening, the person is reckless about whether the act incites hatred toward, revulsion of, serious contempt for, or severe ridicule of, a person or group of people on the ground of disability, gender identity, HIV/AIDS status, intersex status, race, religious conviction and sexuality.*

The penalty is 50 penalty units.

Northern Territories

No civil or criminal anti-vilification laws exist in the Northern Territories.

England and Wales

The principal offences in the United Kingdom which address hate speech are commonly referred to as "stirring up" offences.

The offences are found in the Public Order Act 1986. The Act sets out the offences of stirring up hatred on the grounds of race, religion and sexual orientation.

Part III of the Act sets out the offence of incitement to racial hatred, which was initially created by section 6 of the Race Relations Act 1965.[148] The racial hatred offences apply to specified forms of behavior or content that is:

- threatening, abusive or insulting; and
- intended, or likely, to stir up racial hatred.[149]

This part of the Act applies to England, Scotland and Wales.

[144] Criminal Code 1913, ss 77-80.
[145] Tasmania Anti-Discrimination Act 1998, s 19.
[146] Tasmania Anti-Discrimination Act 1998, s 74.
[147] Discrimination Act 1991, s 67(A).
[148] The 1965 formulation of the Race Relations Act required proof of intention to stir up hatred which is removed from the latest offence.
[149] Race Relations Act, s 18: *A person who uses threatening, abusive or insulting words or behaviour, or displays any written material which is threatening, abusive or insulting, is guilty of an offence if— (a) he intends thereby to stir up racial hatred, or (b) having regard to all the circumstances racial hatred is likely to be stirred up thereby.*

Analogous offences have been created under Part 3A of the Act, which addresses stirring up hatred on the grounds of religion (added in 2017) and sexual orientation (added in 2010).[150] The offences apply to specified forms of behavior or content that is:

- threatening; and
- intended to stir up hatred on these grounds.

These parts do not apply in Scotland. For more on Scotland, see later in this section.

The stirring up hatred on the grounds of religion and sexual orientation offences differ from the provisions for racial hatred in that:

- the behaviour or material must be "threatening", not abusive or insulting;
- the offence can only be committed intentionally (the fact that hatred was likely to be stirred up is insufficient on its own); and
- there are express provisions protecting freedom of expression:

  Section 29J provides:[151]

  Nothing in this Part shall be read or given effect in a way which prohibits or restricts discussion, criticism or expressions of antipathy, dislike, ridicule, insult or abuse of particular religions or the beliefs or practices of their adherents, or of any other belief system or the beliefs or practices of its adherents, or proselytising or urging adherents of a different religion or belief system to cease practising their religion or belief system.

There is a similarly wide protection for the criticism of sexual conduct or practice, and of same sex marriage, in section 29JA:

> *(1) In this Part, for the avoidance of doubt, the discussion or criticism of sexual conduct or practices or the urging of persons to refrain from or modify such conduct or practices shall not be taken of itself to be threatening or intended to stir up hatred.*
>
> *(2) In this Part, for the avoidance of doubt, any discussion or criticism of marriage which concerns the sex of the parties to marriage shall not be taken of itself to be threatening or intended to stir up hatred.*

A range of conduct is caught by the race, religion and sexual orientation based stirring up offences, including displaying, publishing or distributing written material and displaying visual images or sounds.[152]

Racial hatred is defined as hatred against a group of persons by reference to "colour, race, nationality (including citizenship), or ethnic or national origins."[153] Religious hatred is defined as "hatred against a group of persons defined by reference to religious belief or lack of religious belief."[154] Hatred on the grounds of sexual orientation is defined as "hatred against a group of persons defined by reference to sexual orientation (whether towards persons of the same sex, the opposite sex or both)."[155]

[150] The Racial and Religious Hatred Act 2006 came into force on 1 October 2007 and created the new offences of stirring up religious hatred. The Criminal Justice and Immigration Act 2008 amended part 3A of the Public Order Act 1986, which applies to hatred against persons on religious grounds, to create offences of intentionally stirring up hatred on the grounds of sexual orientation. The provisions came into force on 23 March 2010.

[151] Public Order Act 1986, s 29J.

[152] The Public Order Act 1986 includes the following conduct in relation to racial hatred: s 18: *Using threatening, abusive or insulting words or behaviour or displaying written material which is threatening, abusive or insulting; s19: publishing or distributing written material which is threatening, abusive or insulting; s 20: presenting or directing the public performance of a play involving the use of threatening, abusive or insulting words or behaviour; s21: distributing, showing or playing a recording of visual images or sounds which are threatening, abusive or insulting; s 22: providing a programme service, or producing or directing a programme, where the programme involves threatening, abusive or insulting visual images or sounds, or using the offending words or behaviour therein; or s 23: possessing written material, or a recording of visual images or sounds, which is threatening, abusive or insulting, with a view to it being displayed, published, distributed, shown, played or included in a cable programme service.* Similar conduct is caught under ss 29B-29G relating to the grounds of religion or sexual orientation.

[153] Public Order Act 1986, s 17.

[154] Public Order Act 1986, s 29A.

[155] Public Order Act 1986, s 29AB.

On indictment, the maximum penalty for these offenses is seven years imprisonment or a fine or both. On summary conviction, a maximum penalty is six months imprisonment, or a fine not exceeding the statutory maximum, or both.[156]

Potential prosecutions of all incitement crimes must be referred to the Special Crime and Counter Terrorism Division. Cases can only proceed with the consent of the Attorney General.[157]

The Crown Prosecution Service (CPS) plays an important role in the prosecution of hate crimes offences, which include 'hate speech'. The CPS has produced prosecution guidelines and public statements for:

- Racist and Religious Hate Crime
- Stirring up Hatred on the Grounds of Sexual Orientation[158]

According to the CPS, although the success rate in prosecutions is high for stirring up offences, compared to other forms of hate crime, referrals and decisions to charge are much lower.[159] In 2016-17 there were four prosecutions, all with successful outcomes. Between 2008 and 2012, only 113 charges of stirring up racial hatred and 21 charged of stirring up hatred on the ground of religion or sexual orientation reached a first hearing in a magistrate's court.[160]

In contrast 75,000 charges for aggravated offences were laid.[161] These are categories of criminal offences which are aggravated by being accompanied by hostility. There are nine offences which have aggravated versions, including common assault, harassment and stalking. The prosecution must prove not only that the underlying or "basic" offence was committed, but also that in committing it the defendant demonstrated, or was motivated by, hostility.[162] Hateful motivation can also result in enhanced sentencing for other types of offences.

Law Commission 2014 report

In England and Wales, the Law Commission's 2014 report *Hate Crime: Should the Current Offences be Extended?* examined the case for extending the stirring up offences to include disability or transgender identity.[163] The Commission concluded that there was no practical need to extend the offences, stating that: "If new offences of stirring up hatred on the grounds of disability and transgender identity were created, there would be very few successful prosecutions."[164]

The conclusion was based on the following considerations: there have been very few prosecutions for the existing offences of stirring up hatred, the type of hate speech typically found in relation to disability and transgender status is less likely to satisfy the requirements of a stirring up offence than that found in relation to race and religion and they most likely would be covered by other offences; therefore there would be still fewer successful prosecutions for the new stirring up offences than there are now for the existing ones. Therefore, the Commission found that the deterrence and communicative effects of the new offences would be limited.[165]

[156] Public Order Act 1986, s 29L(3).
[157] Public Order Act 1986, ss 27(1) and 29L(1).
[158] Crown Prosecution Service, Homophobic, Biphobic and Transphobic Hate Crimes – Prosecution Guidance (15 August 2018) <www.cps.gov.uk>.
[159] Crown Prosecution Service, Hate Crime Data Report (2017-2018) <www.cps.gov.uk>.
[160] Law Commission, Hate Crime: The Case for Extending the Existing Offences – A Consultation Paper (2014) at [4.8]; Law Commission, Hate Crime: Should the Current Offences be Extended? (May 2014) at 209.
[161] Law Commission, Hate Crime: Should the Current Offences be Extended? (May 2014) at 209.
[162] Crown Prosecution Service, "Hate Crime" (2017) <www.cps.gov.uk/hate-crime>. The UK Police and Crown Prosecution Service (CPS) have agreed on the following definition for identifying and flagging hate crimes: *Any criminal offence which is perceived by the victim or any other person, to be motivated by hostility or prejudice, based on a person's disability or perceived disability; race or perceived race; or religion or perceived religion; or sexual orientation or perceived sexual orientation or transgender identity or perceived transgender identity.*
[163] Law Commission, Hate Crime: Should the Current Offences be Extended? (May 2014) at 209. The initial terms of reference stated disability and gender identity. Gender identity was later confirmed to mean transgender identity.
[164] At 14.
[165] At 209-210.

In 2016, the United Kingdom Home Affairs Parliamentary Select Committee conducted an inquiry into 'Hate crime and its violent consequences.' In May 2017, the Inquiry published a report *"Hate crime: abuse, hate and extremism online."*[166] Among its recommendations, were that:[167]

> *The Government should review the entire legislative framework governing online hate speech, harassment and extremism and ensure that the law is up to date. It is essential that the principles of free speech and open public debate in democracy are maintained—but protecting democracy also means ensuring that some voices are not drowned out by harassment and persecution, by the promotion of violence against particular groups, or by terrorism and extremism.*

The inquiry has been resumed by the current parliament.[168] In September 2018 the Parliamentary Under-Secretary of State for Justice announced she would be asking the Law Commission to complete a wide-ranging review into hate crime to explore how to make current legislation more effective and consider if there should be additional protected characteristics.[169] A consultation report is expected in early 2020.

The first conviction for stirring up hatred on the grounds of sexual orientation occurred in the United Kingdom in 2012.[170] The case involved men who distributed leaflets calling for the death of gay people.[171] The defendants relied on their freedom to preach strongly held beliefs – beliefs they claimed had foundation in scripture.[172] The Court stated that Parliament had not intended to stifle debate but to protect.[173] The Court held that the freedom of expression provision did not extend to the leaflets distributed by the defendants that were found to be threatening to gay people.[174] One defendant was sentenced to two years imprisonment and the other two defendants sentenced to 15 months imprisonment.[175]

Scotland

Scotland only has criminal offences for stirring up hatred in relation to race. The provisions are found in sections 18 to 22 of the Public Order Act 1986, which is the United Kingdom statute that extends to Scotland.

In January 2017, Scotland's Minister for Community Safety and Legal Affairs announced an independent review of hate crime legislation in Scotland to be led by former Privy Council judge, Lord Bracalade. As part of the review, it looked at stirring up hatred offences and recommended in May 2018 that:

- Stirring up hatred offences should be extended beyond racial hatred to other protected characteristics (such as, religion, sexual orientation, transgender identity and disability) including any new protected characteristics.
- Any new stirring up hatred offences should (a) require conduct which is threatening or abusive; and (b) include a requirement (i) of an intention to stir up hatred, or (ii) that having regard to all the circumstances hatred in relation to the particular protected characteristic is likely to be stirred up thereby.
- The current provisions in relation to stirring up racial hatred under the Public Order Act 1986 should be revised and consolidated in a new Act containing all hate crime and stirring up hatred legislation. Any replacement for the stirring up of racial hatred provisions should (a) require conduct

[166] Parliament, *Hate Crime: abuse, hate and extremism online*, (27 April 2017) <www.publications/parliament.uk>.
[167] At [56].
[168] Parliament, Home Affairs Committee: Hate crime and its violent consequences inquiry, (24 April 2017) <www.partliament.uk>.
[169] Law Commission, *Annual Report 2018-2019*, (10 July 2019).
[170] *R v Ali, Javed and Ahmed* (10 Feb 2012) (unreported) in Law Commission, *Hate Crime: The case for extending the existing offence – consultation paper no 213* (2013) at [2.124].
[171] Sentencing remarks of HHJ Burgess in *Ali, Javed and Ahmed* (10 Feb 2012) https://www.judiciary.uk/wp-content/uploads/JCO/Documents/Judgments/sentencing-remarks-r-v-ali-javed-ahmed.pdf
[172] *Ibid.*
[173] *Ibid.*
[174] *Ibid.*
[175] *Ibid.*

which is threatening or abusive; and (b) include a requirement (i) of an intention to stir up hatred, or (ii) that having regard to all the circumstances, hatred in relation to the particular protected characteristic is likely to be stirred up.

- A protection of freedom of expression provision similar to that in sections 29J and 29JA of the Public Order Act 1986 (as outlined above) and section 7 Offensive Behaviour at Football and Threatening Communications (Scotland) Act 2012 (OBFTCA) should be included in any new legislation relating to stirring up offences. Section 7 of the OBFTCA expressly ensures the freedom to debate and express views relating to religion is protected.

The Final Report provides a useful analysis of the merits of "stirring up offences". The Final Report found that such offences recognise wrongfulness, harm, seriousness of the offence and their symbolic function.[176] The Final Report was accompanied by an Academic Report[177] which identified the direct and indirect harm which might result from hate speech:[178]

> *Direct harms are those that might result from members of the targeted community being exposed to hate speech. Indirect harms are those that might result from persons outside the targeted group changing their behaviour or attitudes towards members of the targeted group.*

The Scottish Government carried out public consultations in January and February 2019 to help inform the development of a hate crime bill, following the findings of the Final Report.[179] In the consultation paper the Scottish Government considered that there was merit in considering whether stirring up hatred offences should be extended to other protected characteristics and asked for the publics views on this.[180] It also asked whether the public agreed with Lord Bracalade's recommendation that any new stirring up hatred offences should require that the conduct be 'threatening or abusive' and whether a protection of freedom of expression provision should be included, such as those in the Public Order Act.[181]

An analysis of the 1,159 consultation responses was published on 27 June 2019.[182] The majority of organisations supported the introduction of new offences for all protected characteristics. They emphasised the importance of legal parity and clarity and thought laws of this type could protect all relevant groups while also respecting freedom of speech. In contrast, individuals and faith groups largely disagreed with laws protecting particular groups and/or they had concerns about the impact on freedom of speech and religious expression.[183]

---

[176] Lord Bracadale, *Independent Review of Hate Crime Legislation in Scotland – Final Report* (Scottish Government, Edinburgh, May 2018).

[177] James Chalmers and Fiona Leverick, A Comparative Analysis of Hate Crime (July 2017) <www.consult.gov.scot>. As part of the review, Lord Bracadale requested two academics of the University of Glasgow to produce a comparative report detailing principles underpinning hate crime legislation and approaches taken to hate crime in a range of jurisdictions, known as the academic report.

[178] At 69.

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

[180] *Ibid*.

[181] At 43. Public Order Act s 29J "Protection of freedom of expression" states *Nothing in this Part shall be read or given effect in a way which prohibits or restricts discussion, criticism or expressions of antipathy, dislike, ridicule, insult or abuse of particular religions or the beliefs or practices of their adherents, or of any other belief system or the beliefs or practices of its adherents, or proselytising or urging adherents of a different religion or belief system to cease practising their religion or belief system*. Section 29JA Protection of freedom of expression (sexual orientation) states: (1) *In this Part, for the avoidance of doubt, the discussion or criticism of sexual conduct or practices or the urging of persons to refrain from or modify such conduct or practices shall not be taken of itself to be threatening or intended to stir up hatred. (2) In this Part, for the avoidance of doubt, any discussion or criticism of marriage which concerns the sex of the parties to marriage shall not be taken of itself to be threatening or intended to stir up hatred*. Lord Bracalade also referred to s 7 of the Offensive Behavior at Football and Threatening Communications (Scotland) Act 2012 - Protection of freedom of expression which states: (1)*For the avoidance of doubt, nothing in section 6(5) prohibits or restricts—(a)discussion or criticism of religions or the beliefs or practices of adherents of religions, (b)expressions of antipathy, dislike, ridicule, insult or abuse towards those matters, (c)proselytising, or (d)urging of adherents of religions to cease practising their religions. (2) In subsection (1), "religions" includes— (a)religions generally, (b)particular religions, (c)other belief systems*.

[182] Scottish Government, One Scotland: Hate has no home here – Consultation on amending Scottish hate crime legislation Analysis of responses Final Report (June 2019) at [2].

[183] At [19].

There were differing views on how any new offences should be formulated – while some were happy with the wording proposed in the consultation, others were concerned the wording would set the threshold too high or, conversely, too low. However, there was broad agreement that stirring up offences should include a protection of freedom of expression provision, although there were differing views on how this should be framed.[184]

Northern Ireland

The Public Order (Northern Ireland) Act 1987 includes the offences of acts intended or likely to stir up hatred or arouse fear, which corresponds closely to the principal relevant offences under the Public Order Act 1986. The offence of incitement to hatred defines "hatred" as "hatred against a group of persons in the State or elsewhere on account of their race, colour, nationality (including citizenship), religion, ethnic or national origins, disability or sexual orientation".[185] The offences cover the use of words or behaviour or display of written material,[186] publishing or distributing written material,[187] distributing, showing or playing a recording,[188] broadcasting or including programme in cable programme service,[189] and require that the individual:

(1) Intends to stir up hatred or arouse fear; or

(2) Having regard to all the circumstances hatred is likely to be stirred up or fear is likely to be aroused thereby.

Ireland

The Prohibition of Incitement to Hatred Act 1989 makes it an offence to:

(a) to publish or distribute written material,

(b) to use words, behave or display written material—

(i) in any place other than inside a private residence, or

(ii) inside a private residence so that the words, behaviour or material are heard or seen by persons outside the residence, or

(c) to distribute, show or play a recording of visual images or sounds,

if the written material, words, behaviour, visual images or sounds, as the case may be, are threatening, abusive or insulting and are intended or, having regard to all the circumstances, are likely to stir up hatred.

The offence of incitement of hatred defines "hatred" as "hatred against a group of persons in the State or elsewhere on account of their race, colour, nationality, religion, ethnic or national origins, membership of the travelling community or sexual orientation."[190]

Canada

Canada has both civil and criminal hate speech laws. Canada's criminal hate speech laws are found at the federal, rather than provincial level, and are included in the Canadian Criminal Code. Following the 2013 repeal of the civil hate speech provisions of the Canadian Human Rights Act[191] there are no longer any civil hate speech laws at federal level. However, some form of civil hate speech laws exist in most provinces.

[184] At [20].
[185] Public Order (Northern Ireland) 1987, s 8.
[186] Section 9.
[187] Section 10.
[188] Section 11.
[189] Section 12.
[190] Prohibition of Incitement to Hatred Act 1989, s 1.
[191] Section 13 of the Canadian Human Rights Act 1985 made it a discriminatory practice to communicate by telephone or the internet that is "likely to expose a person or persons to hatred or contempt by reason of the face that the person or those persons are identifiable on the basis of a prohibited ground of discrimination." The Canadian Human Rights Commission (CHRC) published two reports in 2008 and 2009 concerning section 13. The first, released in October 2008, was written by Richard Moon, a Canadian law professor. Professor Moon recommended that s 13 of the CHRA should be repealed so that the Canadian Human Rights Commission and the Canadian Human Rights Tribunal no longer deal with hate speech, and in particular hate speech on the Internet. The CHRC produced a Special Report to Parliament in 2009 and concluded that both the Criminal Code and the CHRA serve valid purposes in dealing with hate messages on the Internet. It therefore did not support the repeal of s 13; however, it proposed a number of reforms. Bill C-304 was introduced in the House of Commons by a Member of Parliament on 30 September 2011. In 2012 the House of Commons voted to repeal s 13(1) and on 26 June 2013 it was finally repealed.

The Canadian Criminal Code contains three types of hate speech offences in a section labelled "hate propaganda". These are advocating genocide; public incitement of hatred, and wilful promotion of hatred. These hate propaganda offences were added to the Canadian Criminal Code in 1970 in response to events and developments in the 1960s when white supremacists and neo-Nazi groups were active in Canada.[192]

Section 318(1) – *incitement of genocide* – provides that "Everyone who advocates or promotes genocide is guilty of an indictable offence and liable to imprisonment for a term not exceeding five years." No prosecution under this provision can be undertaken without the consent of the provincial Attorney General.[193]

Section 319(1) – *public incitement of hatred* – provides that "Everyone who, by communicating statements in any public place, incites hatred against any identifiable group where such incitement is likely to lead to a breach of the peace" is guilty of a criminal offence.

Section 319(2) – *wilful promotion of hatred* – provides "Everyone who, by communicating statements, other than in private conversation, wilfully promotes hatred against any identifiable group" is guilty of a criminal offence.

Any person charged under section 319(2) has four special defences available.[194] These are:

- that the statements communicated were true;
- that an opinion or argument was expressed in good faith and either concerned a religious subject or was based on a belief in a religious text;
- that the statements were relevant to a subject of public interest and were on reasonable grounds believed to be true; and
- that the statements were meant to point out matters that produce feelings of hatred toward an identifiable group and were made in good faith for the purpose of their removal.

Both of the section 319 offences carry a maximum penalty of imprisonment of two years when prosecuted on indictment.[195] No prosecution under these provisions can be undertaken without the consent of the provincial Attorney General.[196]

Sections 319(1) and 319(2) both refer to an identifiable group, which means "any section of the public distinguished by colour, race, religion, national or ethnic origin, age, sex, sexual orientation, or mental or physical disability."[197] "Communicating" includes "communicating by telephone, broadcasting or other audible or visible means" and "statements" include "words spoken or written or recorded electronically or electro-magnetically or otherwise, and gestures, signs or other visible representations".[198]

The main difference between the two offences is that section 319(1) requires that the accused's conduct is likely to result in some sort of public disorder (a breach of the peace), whereas section 319(2) requires only that the accused willfully promoted hatred.

In *R v Keegstra* the Canadian Supreme Court examined the constitutionality of the section 319(2) offence of wilful promotion of hatred. Dickson CJ, in the majority judgment found that it did impinge on freedom of expression as protected by section 20(b) of the Canadian Charter of Rights and Freedoms.[199] The restriction, however, was justified under section 1 of the Charter (limitation provisions) because it served the important purpose of preventing the spreading of hate propaganda[200] and advanced this goal rationally and with minimal impairment of freedom of expression.[201]

[192] Julian Walker, "Background Paper: Hate Speech and the Freedom of Expression: Legal Boundaries in Canada", Library of Parliament, Publication No 2018-25-E (29 June 2018) Paper at 4.
[193] Canadian Criminal Code, s 318(3).
[194] Set out in s 319(3).
[195] Canadian Criminal Code, s 319(1)(a) (public incitement of hatred); s 319(2)(a) (wilful promotion of hatred).
[196] Section 318(3).
[197] Section 318(4).
[198] Section 319(7).
[199] *R v Keegstra* [1990] 3 SCR 697, per Dickson CJ at 730 (emphasis added).
[200] At795.
[201] At 766.

In relation to the second point (the extent of the limitation placed on freedom of expression by section 319(2)), it was relevant that the mental element of the section 319(2) offence (wilful promotion) is narrow in scope and captures "only the most intentionally extreme forms of expression."[202] As such, the limitations placed on freedom of expression by section 319(2) are minimal.

Dickson CJ also noted:[203]

> *hate propaganda opposes the targeted group's ability to find self-fulfillment by articulating their thoughts and ideas. It impacts on that group's ability to respond to the substantive ideas under debate, thereby placing a serious barrier to their full participation in our democracy. Indeed, a particularly insidious aspect of hate speech is that it acts to cut off any path of reply by the group under attack.* ***It does this not only by attempting to marginalize the group so that their reply will be ignored: it also forces the group to argue for their basic humanity or social standing, as a precondition to participating in the deliberative aspects of our democracy.***

Canada's 10 territories also have human rights legislation that provides civil remedies. Most of the legislation applies to an expansive range of protected characteristics such as.[204]

- Race
- Religious belief
- Colour
- Gender
- Gender identity
- Gender expression
- Physical disability
- Mental disability
- Age
- Ancestry
- Place of origin
- Marital status
- Source of income
- Family status
- Sexual orientation

In 2013, the Supreme Court of Canada affirmed the legitimacy of human rights legislation that restricts hate speech in *Sakatchewan (Human Rights Commission) v Whatcott*.

In *Whatcott*, the Supreme Court was tasked with determining whether the hate speech provision under section 14(1)(b) of the Saskatchewan Human Rights Code is constitutional in light of the right to freedom of expression. The case concerned four complaints that were filed with the Saskatchewan Human Rights Commission about four flyers allegedly promoting hatred against individuals on the basis of their sexual orientation that were published and distributed by Mr Whatcott.

In a unanimous decision, the Supreme Court held that the Saskatchewan hate speech provision (with some words removed because they were overbroad) was justified and constitutional.

The Court considered the definitions of "hatred and contempt," affirming the analysis of the Court in 1990 in *R v Keegstra* and *Taylor v The Canadian Human Rights Commission*, and setting out the approach that courts and tribunals should use in interpreting these terms. The Court explained that they refer to expression of an unusual and extreme nature, involving vilification, dehumanization, and reviling. This interpretation excludes merely offensive expression.

The Court noted that freedom of expression is central to democracy, but it is not absolute, and limitations may be justified under section 1 of the Charter. In considering justification for the limit, the Court held that the objective of the legislation is pressing and substantial.

[202] At 783.
[203] At 763.
[204] Alberta, British Columbia, Prince Edward Island, Manitoba, New Brunswick, New Foundland, Nova Scotia, Ontario, British Columbia, Quebec, Saskatchewan.

The Court recognised the harm caused by hate speech, not only to the targeted group, but also to society at large. "Hate speech lays the groundwork for later, broad attacks on vulnerable groups," Justice Rothstein wrote. "These attacks can range from discrimination, to ostracism, segregation, deportation, violence and, in the most extreme cases, to genocide." A "particularly insidious" effect of hate speech is that it inhibits the expression of the targeted group, Justice Rothstein noted.[205]

The Court went on to hold that the prohibition against hate speech involves balancing between freedom of expression and equality rights. The Court concluded that the limitation on freedom of expression by the prohibition of hate speech, when properly defined and understood, is demonstrably justified in a free and democratic society.

United States

The First Amendment to the United States Constitution provides substantial protection for speech no matter how offensive its content.[206] However, there are categories of speech that receive no First Amendment protection. These are fighting words, libel, obscenity, child pornography and true threats.

The First Amendment does not protect conduct that crosses the line into targeted harassment or threats, or that creates a pervasively hostile environment.[207]

Hate crimes

There are no statutes prohibiting hate speech in the United States, however there are federal statutes against hate crimes. Federal criminal civil rights laws impose criminal penalties for the deprivation of certain federal rights, privileges, or immunities. A majority of these laws prohibit violent and intimidating acts motivated by animus based on race, ethnicity, national origin, religious beliefs, gender, sexual orientation, or disability.[208]

The United States Department of Justice enforces federal hate crimes laws that cover certain crimes committed on the basis of race, colour, religion, national origin, sexual orientation, gender, gender identity, or disability. The Department of Justice began prosecuting federal hate crimes cases after the enactment of the Civil Rights Act of 1968.[209] The 1968 statute made it a crime to use, or threaten to use, force to wilfully interfere with any person because of race, colour, religion, or national origin and because the person is participating in a federally protected activity, such as public education, employment, jury service, travel, or the enjoyment of public accommodations, or helping another person to do so.[210]

In 1968, United States Congress also made it a crime to use, or threaten to use, force to interfere with housing rights because of the victim's race, colour, religion, sex, or national origin.[211] in 1988, protections on the basis of familial status and disability were added.

In 1996, Congress passed the Church Arson Prevention Act.[212] Under this Act, it is a crime to intentionally deface, damage, or destroy religious real property, or interfere with a person's religious practice, in situations affecting interstate or foreign commerce.[213] The Act also bars intentionally defacing, damaging, or destroying religious property because of the race, colour, or ethnicity of persons associated with the property.

[205] *Sakatchewan (Human Rights Commission) v Whatcott* [2013] 1 SCR 467 at [75].
[206] United States Constitution, amendment I.
[207] American Library Association "Hate Speech and Hate Crime" (December 2017) <www.ala.org/advocacy/intfreedom/hate>.
[208] Alison M. Smith, "*Overview of Selected Federal Criminal Civil Rights Statutes*" Congressional Research Service Report Prepared for Members and Committees of Congress (16 December 2014) < https://fas.org/sgp/crs/misc/R43830.pdf > at 1.
[209] The United States Department of Justice "Federal Laws and Statutes" < https://www.justice.gov/hatecrimes/laws-and-policies>.
[210] Civil Rights Act of 1968 Pub Law No 90–284, § 82, Stat 73.
[211] Criminal Interference with Right to Fair Housing 42 USC § 3631.
[212] Damage to Religious Property, Church Arson Prevention Act 18 USC § 247.
[213] Damage to Religious Property, Church Arson Prevention Act 18 USC § 247.

The Violent Interference with Federally Protected Rights legislation makes it a crime to use or threaten to use force to wilfully interfere with a person's participation in a federally protected activity because of race, colour, religion, or national origin. Federally protected activities include public education, employment, jury service, travel, or the enjoyment of public accommodations. Under this statute, it is also a crime to use or threaten to use force against those who are assisting and supporting others in participating in these federally protected activities.[214]

The Matthew Shepard and James Byrd, Jr. Hate Crimes Prevention Act of 2009 was the first statute allowing federal criminal prosecution of hate crimes motivated by the victim's actual or perceived sexual orientation or gender identity.[215] The Act makes it a federal crime to wilfully cause bodily injury, or attempt to do so using a dangerous weapon, because of the victim's actual or perceived race, colour, religion, or national origin. The Act also covers crimes committed because of the actual or perceived religion, national origin, sexual orientation, gender, gender identity, or disability of any person, if the crime affected interstate or foreign commerce or occurred within federal special maritime or territorial jurisdiction.

The Conspiracy Against Rights statute makes it unlawful for two or more persons to conspire to injure, threaten, or intimidate a person in any state, territory, or district in the free exercise or enjoyment of any right or privilege secured to him or her by the Constitution or the laws of the United States.[216] Punishment for violations includes a fine and/or imprisonment for a maximum of 10 years. The law provides greater punishment to violators if their acts result in death (or an attempt to kill) or include kidnapping (or an attempt to kidnap) or aggravated sexual abuse (or an attempt to commit aggravated sexual abuse). Under such heightened circumstances, offenders may face life imprisonment or the death penalty.[217]

The Deprivation of Rights Under Color of Law legislation makes it a crime for "any person acting under color of any law, statute, ordinance regulation, or custom to willfully deprive or cause to be deprived from any person those rights, privileges, or immunities secured or protected by the Constitution and laws of the U.S."[218] Federal civil rights prosecutions against state actors (e.g., law enforcement) are usually conducted pursuant to this statute.

Many American states specifically make provision for sentence aggravation and collection of hate crime data.[219] Most states and United States territories have hate crime statutes that are enforced by state and local law enforcement in state and local courts. Hate crime laws in states and territories vary widely across jurisdictions:[220]

- Bias motivations: Different jurisdictions define hate crimes to include different bias motivations.
- Penalty enhancements: Laws in some jurisdictions increase the sentence for crimes motivated by identified factors. At least 46 states and the District of Columbia have statutes with penalties for bias-motivated crimes.
- Data collection: Some jurisdictions require collecting data on hate crimes. Data provides better transparency into crimes that are occurring and can help states allocate support and resources to communities in greatest need.

Even if a state or territory does not have a hate crimes law, hate crimes can still be reported to the Federal Bureau of Investigation.

[214] Violent Interference with Federally Protected Rights 18 USC § 245.
[215] The Matthew Shepard and James Byrd, Jr. Hate Crimes Prevention Act of 2009 18 USC § 249.
[216] Conspiracy Against Rights 18 USC § 241.
[217] Alison M. Smith, "Overview of Selected Federal Criminal Civil Rights Statutes", above n 9, at 2.
[218] Deprivation of Rights Under Color of Law 18 USC §242.
[219] Anti-Defamation League Washington Office, Anti-Defamation League State Hate Crime Statutory Provisions, (April 2018), <www.adl.org>.
[220] The United States Department of Justice "Federal Laws and Statutes", above n 10.

Hate speech

However, regulation of hate speech on the other hand is very limited given the constitutional protection of freedom of speech. Under current First Amendment jurisprudence, speech can only be regulated when it directly incites imminent criminal activity or consists of specific threats of violence targeted against a person or group.

In *Beauharnais v Illinois* (1952), Justice Felix Frankfurter held that a leaflet accusing black people as a group of being rapists, robbers, carriers of guns and knives, and drug users was libel. He elaborated that libel against a group, like libel against an individual, is not within the area of constitutionally protected speech.[221] Justice Frankfurter outlined instances where speech may be curtailed, citing directly from the Supreme Court's decision in *Chaplinksy v New Hampshire*, including the "lewd and obscene, the profane, the libelous and the insulting or 'fighting' words — those which, by their very utterance, inflict injury or tend to incite an immediate breach of the peace."[222]

Although, *Beauharnais* has not been overturned, it must be read in conjunction with *Collin v Smith* and *Brandenburg v Ohio*, below. Following these decisions, if the speech does not produce imminent harm, then it is likely to be regarded as protected speech.

In *Brandenburg v Ohio* (1969), the Supreme Court protected a Ku Klux Klan member's hateful and disparaging speech directed towards African-Americans, holding that such speech could only be limited if it posed an "imminent danger" of inciting violence. The Supreme Court ruled that a state could only forbid or proscribe advocacy that is "directed to inciting imminent lawless action and is likely to incite or produce such action."[223] The standard in *Brandenburg* makes it immensely difficult to justify restrictions on hate speech and is very speech protective: all speech is protected, except speech that is intentionally directed at, and likely to cause, lawless action.

In *Collin v Smith* (1978), the Seventh Circuit Court of Appeals upheld a decision that allowed a group of neo-Nazis to march on the streets of an Illinois suburb housing a substantial Jewish population that included Holocaust survivors. The court stated that, "above all else, the First Amendment means that government has no power to restrict expression because of its message, its ideas, its subject matter, or its content."[224] It further stated that, "if these civil rights are to remain vital for all, they must protect not only those society deems acceptable, but also those whose ideas it quite justifiably rejects and despises".[225]

In *R.A.V. v City of St. Paul* (1992), the Supreme Court overturned the conviction of a teenager convicted of burning a cross on the lawn of an African American family's home.[226] The Court held that a few limited categories of speech, such as obscenity, defamation, and fighting words, may be regulated because of their constitutionally proscribable content. However, these categories are not entirely invisible to the Constitution, and government may not regulate them based on hostility, or favouritism, towards a nonproscribable message they contain.[227]

This is based upon the belief that freedom of speech requires the government to strictly protect robust debate on matters of public concern even when such debate devolves into distasteful, offensive, or hateful speech that causes others to feel grief, anger, or fear.

[221] *Beauharnais v Illinois* 343 US 250 (1952).
[222] *Chaplinsky v New Hampshire* 315 US 568 (1942) at 571-572 as cited in *Beauharnais v Illinois* 343 US 250 (1952) at 256.
[223] *Brandenburg v Ohio* 395 US 444 (1969) at 448.
[224] *Collin v Smith* 578 F 2d 1197 (7th Cir 1978) at 1202.
[225] At 1210.
[226] *R.A.V. v City of St. Paul* 505 US 377 (1992).
[227] At 382-390.

More recently, the Supreme Court has acknowledged the offensive nature of hate speech in cases like *Matal v Tam* (2017), but they have been reluctant to impose broad restrictions on it:[228]

> *Speech that demeans on the basis of race, ethnicity, gender, religion, age, disability, or any other similar ground is hateful; but the proudest boast of our free speech jurisprudence is that we protect the freedom to express the thought that we hate.*

Matal v Tam concerned a dance-rock band's application for federal trademark registration of the band's name, "The Slants." "Slants" is a derogatory term for persons of Asian descent, and members of the band are Asian-Americans. But the band members believed that by taking that slur as the name of their group, they would help to "reclaim" the term and drain its denigrating force.[229] The Supreme Court confirmed by a unanimous decision that prohibiting the registration of trademarks that may 'disparage' persons, institutions, beliefs, or national symbols with the United States Patent and Trademark Office violated the First Amendment.

Justice Kennedy stated, "a law that can be directed against speech found offensive to some portion of the public can be turned against minority and dissenting views to the detriment of all. The First Amendment does not entrust that power to the government's benevolence. Instead, our reliance must be on the substantial safeguards of free and open discussion in a democratic society."[230]

[228] *Matal v Tam* 582 US __ (2017) at 25.
[229] At 1.
[230] At 38.

Part V: Concluding Comments

This paper has provided an overview of hate speech in the New Zealand regulatory context and a summary of the different approaches taken in other countries as well as the relevant international human rights requirements. The document is intended to assist understanding of the important rights and freedoms that are pertinent to any discussion about hate speech and to contribute to the current public discourse and debate. It is hoped that it will assist ongoing and respectful dialogue about our current legal framework.

Appendix 1

Human Rights Act 1993

Section 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.

(2) It shall not be a breach of subsection (1) to publish in a newspaper, magazine, or periodical or broadcast by means of radio or television or other electronic communication a report relating to the publication or distribution of matter by any person or the broadcast or use of words by any person, if the report of the matter or words accurately conveys the intention of the person who published or distributed the matter or broadcast or used the words.

(3) For the purposes of this section,—

newspaper means a paper containing public news or observations on public news, or consisting wholly or mainly of advertisements, being a newspaper that is published periodically at intervals not exceeding 3 months

publishes or distributes means publishes or distributes to the public at large or to any member or members of the public

written matter includes any writing, sign, visible representation, or sound recording

Section 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.

(2) For the purposes of this section, publishes or distributes and written matter have the meaning given to them in section 61.

The table below sets out the protected characteristics in relation to criminal laws:

Criminal hate speech laws

	NZ	Aus									Canada	England & Wales	Ireland	North Ireland	Scotland	United States*
		Fed	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	Fed					
Race	●		●	●	●	●	●			●	●	●	●	●	●	
National origin	●										●	●	●	●	●	
Ethnic origin	●										●	●	●	●	●	
Colour	●										●	●	●	●	●	
Religion			●	●	●					●	●	●	●	●		
Sex (gender)											●					
Sexual orientation			●		●					●	●	●	●	●		
Gender expression											●					
Gender identity			●		●					●	●					
Intersex			●							●						
HIV status			●							●						
Age											●					
Disability										●	●			●		
Member of travelling community													●			

*The United States does not criminalise any hate speech based on particular characteristics. Rather there are specific categories of unprotected speech: these are fighting words, libel, obscenity, child pornography and true threats. Speech can only be regulated when it directly incites imminent criminal activity or consists of specific threats of violence targeted against a person or group — the court does not allow special protection for any particular groups. However, speech cannot be restricted because of its message, its ideas, its subject matter, or its content.

NZ Human Rights.

Human Rights Commission
Te Kāhui Tika Tangata