

ARTICLE

When Silence Isn't Golden: How the Silencing Effect of Speech Provides Justification for Hate Speech Regulation

KATHERINE WERRY*

Despite substantial literature on hate speech and its harmful effects, theorists have struggled to justify regulating hate speech and limiting freedom of speech. This article argues that the silencing effect of hate speech provides sufficient basis to justify its regulation and counter free speech objectors in New Zealand. Hate speech—be it spoken or written, words or images—has the capability to silence its victims by limiting their opportunities for speech. In this article, I uncover the inherent contradiction in arguing that regulating hate speech cuts across freedom of speech on the basis that unregulated hate speech in fact limits the speech of some people or groups. The concept that speech may operate to silence is sourced from feminist literature exploring the silencing effect that pornography has on women. This literature can be credited with developing a cohesive definition of silencing and furthering the existing work that examines the harm caused by hate speech. I draw from the feminist line of argument and apply it to the context of racist hate speech and racial minorities. I make three principal claims in this article: hate speech can silence, hate speech does silence and the silencing effect of hate speech means that there should be greater regulation of it in New Zealand. This article is a call for action against the historic and systemic silencing of racial groups, which is still widespread today. Legislative reform of New Zealand's hate speech laws is one way of giving opportunities for speech back to those who have been deliberately silenced.

* BA/LLB(Hons), University of Auckland. I would like to thank my supervisor, Dr Jane Norton, for her constructive feedback and guidance. I would like to acknowledge the work of Catharine MacKinnon, which initially sparked my interest in silencing. I would also take this opportunity to thank my mother, who read numerous drafts of this article without complaint. Your support has been invaluable.

I Introduction

[T]here's really no such thing as the "voiceless". There are only the deliberately silenced, or the preferably unheard.

—Arundhati Roy¹

Hate speech—just saying those two words provokes controversy. The ongoing debates around hate speech and freedom of speech seem never-ending. Many studies have shown that hate speech results in harm,² yet any attempts to regulate it are met with inevitable backlash.³ The difficulty lies in finding a justification for limiting freedom of speech. Without such justification, criticisms of hate speech are unsubstantiated.

In this article, I argue that hate speech in New Zealand should be subject to stronger regulations. The justification for this is that hate speech operates to silence racial minorities. Not all silence is golden—the silencing of racial groups can have real and harmful effects. This silencing argument is based on the feminist arguments—advanced by theorists such as Catharine MacKinnon and Rae Langton—that pornography silences women.⁴

Part II summarises the feminist argument on how speech can silence groups in society, and in particular, how pornography can silence women. Part III addresses how this feminist argument is transferable to race. While neither pornography and hate speech nor sex and race are exactly equivalent, I conclude that this does not preclude the silencing argument from applying to hate speech. Part IV applies the silencing argument to racial hate speech. I define silencing and show how hate speech can silence in each of the three different categories of silencing. I then cover additional elements of the silencing argument such as the prevention of public participation, drawing on Jeremy Waldron's work,⁵ and the interaction between power and speech.

In Part V, I outline and rebut four likely objections to the argument that hate speech silences racial minorities: that hate speech does not literally silence, that strengthening free speech strengthens it for all people (including minorities), that the state should not act as a censor and that offensive speech is still entitled to protection. Finally, in Part VI, I argue that the silencing argument can provide a sufficient justification for laws regulating hate speech in New Zealand.

-
- 1 Arundhati Roy "Peace & The New Corporate Liberation Theology" (City of Sydney Peace Prize Lecture, Sydney, 3 November 2004).
 - 2 For an overview of various studies, see Mari J Matsuda "Public Response to Racist Speech: Considering the Victim's Story" (1989) 87 Mich L Rev 2320 at 2335–2341.
 - 3 See, for example, Stuart Dye "Backlash on hate speech proposal" *The New Zealand Herald* (online ed, 18 March 2005).
 - 4 See, for example, Catharine A MacKinnon *Only Words* (Harvard University Press, Cambridge (Mass), 1993); and Rae Langton "Speech Acts and Unspeakable Acts" (1993) 22 *Philosophy & Public Affairs* 293 at 297–298.
 - 5 Jeremy Waldron *The Harm in Hate Speech* (Harvard University Press, Cambridge (Mass), 2012).

II Feminist Argument: Pornography Silences Women

Feminist theorists argue that pornography should be regulated because pornography and its protection has silenced and deprived women of their speech.⁶ The first step in their argument is to show that speech in general can operate to silence. This claim is most useful for the purposes of my article. In this Part, I examine the general claim that speech can have a silencing effect, and the more specific feminist argument that pornography silences women. I also briefly cover the claim that silencing has harmful effects.

A *General silencing*

The central claim of the silencing argument is that some speech may silence by preventing others from exercising their freedom of speech.⁷ Speech can constrain the actions and speech of those it is directed towards.⁸ It can remove opportunities to speak and be heard—opportunities that are available to other people in society.

Silencing speech often occurs in everyday life. For example, authority figures giving orders are often given specific power to silence (picture a judge saying “silence in the court”).⁹ Authority figures have usually been given the power to silence through orders with the understanding that this will prevent harm or achieve some other beneficial purpose. The judge saying “silence in the court” is restoring order when proceedings become disruptive. The Speaker of the House in Parliament has a similar role. Speech in this context, however, is not the focus of this article. Speech outside of these contexts can also silence, and this is where potential harm arises.

Powerful people, by virtue of their relative status in society, can silence the speech of powerless groups.¹⁰ These groups include women (as per the feminist argument) and racial groups—the focus of this article.

B *Silencing effect of pornography*

The feminist argument focuses on how pornography, in particular, silences women. MacKinnon claims that “pornography amounts to terrorism and promotes not freedom but silence”.¹¹ Frank I Michelman demonstrates the myriad of ways in which pornography can silence women. It prevents women being able to respond effectively with speech of their own, and silences women’s speech by using images and themes in pornography that change consumers’ perceptions of women.¹²

6 MacKinnon, above n 4, at 9. Note that the view of pornography today may differ from the following feminist theory. Women frequently now access and consume pornography. They may also have more control over the creation of pornography.

7 Ronald Dworkin “Liberty and Pornography” *The New York Review of Books* (New York, 15 August 1991) at 15.

8 Langton, above n 4, at 328.

9 At 318.

10 At 299.

11 Catharine A MacKinnon *Feminism Unmodified: Discourses on Life and Law* (Harvard University Press, Cambridge (Mass), 1987) at 130.

12 Frank I Michelman “Alumni Distinguished Lecture in Jurisprudence: Conceptions of Democracy in American Constitutional Argument: The Case of Pornography Regulation” (1989) 56 *Tenn L Rev* 291 at 299–301.

Silencing may occur through the objectification of women. Consumers of pornography are inundated with content that portrays women as objects for male gratification.¹³ This objectification strips women's speech of credibility because "objects do not speak".¹⁴ The most prevalent area in which women's speech is stripped of credibility is in relation to accounts of sexual assault.¹⁵ The silencing of rape victims is one of the most harmful effects of pornography that feminist theorists focus on.

C *Effects of silencing*

Feminist theorists emphasise the damaging effects of silencing. Silencing is a form of subordination, and deprives people of power and liberties that the rest of the populace has.¹⁶ A form of speech such as pornography discredits its targets, reduces their authority and invalidates their speech.¹⁷ It perpetuates the narrative, already prevalent in society, that women's words, thoughts and experiences are not of the same value as those of men. Throughout history, women have been told that they are stupid—that their thoughts are trivial, and their experiences unspeakable.¹⁸ Pornography reinforces these harmful messages by being a substitute for women's lived experiences, and "constructing the social reality of what a woman is".¹⁹

I discuss the silencing argument to a greater extent in Part IV, where I apply the argument to hate speech. Before I proceed to this, it is necessary to compare hate speech against pornography to explore the variations in the silencing argument for these two types of speech.

III Hate Speech versus Pornography

As the silencing argument was first raised in relation to pornography, there may be difficulties in applying the argument to hate speech. This Part considers these problems and concludes that they do not detract from the strength of the assertion that hate speech silences racial groups. I define hate speech, then compare hate speech to pornography, and finish by comparing sex and race. Differences that arise do not take away from the fundamental premise that both pornography and hate speech can operate to silence.

A *Definition of hate speech*

Hate speech can be defined in a multitude of ways, and it is difficult to come up with a cohesive definition. Jacinda Ardern, in 2019, said that "when you see it, you know it"²⁰—echoing Justice Stewart's well-known comment on obscenity.²¹ In literature, rigid definitions of hate speech are often avoided and scholars tend to provide instances of hate

13 MacKinnon, above n 11, at 182.

14 At 182.

15 At 193; and see, for example, Susan Brownmiller *Against Our Will: Men, Women and Rape* (Bantam Books, New York, 1976) at 394.

16 Langton, above n 4, at 329.

17 MacKinnon, above n 11, at 193.

18 At 56–57.

19 MacKinnon, above n 4, at 25.

20 Interview with Jacinda Ardern, Prime Minister (Duncan Garner, The AM Show, Three, 30 April 2019) sourced from Facebook <<https://fb.watch/1jiekwa7mm/>> at 00:03:00–00:03:02.

21 *Jacobellis v State of Ohio* 378 US 184 (1964) at 197.

speech instead of definitions.²² However, many scholars have attempted to categorise or define hate speech, and from these, common themes emerge.²³ There are three main elements that most theorists agree are common to hate speech: the content or message of the speech itself, the intention behind the speech and the effects of the speech.

I rely on the definition of Alice Marwick and Ross Miller, which is based on a comprehensive review of existing hate speech literature.²⁴ I have not used the narrow legislative definition, as this article calls for greater regulation of hate speech in the law, which may include amending or improving that definition. Marwick and Miller define hate speech as having three general elements: content-based, intent-based and harm-based.²⁵ The content-based element refers to the words or images used in the speech, which should be considered as both “offensive to a particular minority and objectively offensive to society”.²⁶ The intent-based element refers to the speaker’s sole intention to “promote hatred, violence or resentment against a particular minority” or an individual member of a minority group.²⁷ Finally, the speech must cause the victim harm. This is not limited to physical harm. Calvin R Massey gives examples of the possible types of harm as follows: “loss of self-esteem, economic and social subordination, physical and mental stress, silencing of the victim, and effective exclusion from the political arena.”²⁸ Hate speech can lead to physiological symptoms and emotional distress, post-traumatic stress disorder, psychosis and suicide.²⁹ Furthermore, harm is to be assessed subjectively, not objectively.³⁰

This article focuses on hate speech on the basis of race only. This is not to discredit the harm that hate speech causes to other minority groups. I have narrowed my focus for the following reasons: first, to mitigate against the risk of diluting the strength of my argument by having too wide of a scope. Second, because hate speech directed towards racial groups is already recognised under New Zealand law, albeit imperfectly, as I will discuss further in Part VI. Third, because racial hate speech is pervasive and recognisable in New Zealand; a significant proportion of hate speech is directed towards Māori and Pasifika people.³¹ Finally, given the Black Lives Matter movement, hate speech towards racial groups is of

22 Andrew F Sellars *Defining Hate Speech* (Berkman Klein Center for Internet & Society, Research Publication No 2016-20, December 2016) at 14–15.

23 For example, Richard Delgado proposed a three-part test for racist speech: Richard Delgado “Words that Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling” (1982) 17 Harv CRCL L Rev 133 at 179. Mari Matsuda defines three identifying characteristics of racist hate messages as being of racial inferiority, directed towards a historically oppressed group, and prosecutorial, hateful and degrading: Matsuda, above n 2, at 2357. Mayo Moran has a wider definition of hate speech, defining it as “speech that is intended to promote hatred against traditionally disadvantaged groups”: Mayo Moran “Talking About Hate Speech: A Rhetorical Analysis of American and Canadian Approaches to the Regulation of Hate Speech” [1994] Wis L Rev 1425 at 1430.

24 Alice Marwick and Ross Miller *Online Harassment, Defamation, and Hateful Speech: A Primer of the Legal Landscape* (Fordham Center on Law and Information Policy, 10 June 2014).

25 At 16.

26 At 16.

27 At 17.

28 Calvin R Massey “Hate Speech, Cultural Diversity, and the Foundational Paradigms of Free Expression” (1992) 40 UCLA L Rev 103 at 105, n 2.

29 Matsuda, above n 2, at 2336.

30 Marwick and Miller, above n 24, at 17.

31 For example, see *Wall v Fairfax New Zealand Ltd* [2018] NZHC 104, [2018] 2 NZLR 471; and Mandy Te “Racism complaints to Human Rights Commission hit five year high” (22 February 2018) Stuff <www.stuff.co.nz>.

great topical significance. I hope that this article will prove useful for other minority groups to utilise the silencing argument in the future.

B *Comparing hate speech and pornography*

The key problem with applying the feminist silencing argument to hate speech is that pornography and hate speech are often classified as different types of speech that are entitled to different levels of government protection.

Jurisprudence in the United States has drawn a distinction between two types of speech: low-value and high-value speech.³² There is no clear-cut definition as to what constitutes either category of speech. However, Cass R Sunstein summarises low-value speech as non-political and non-cognitive material.³³ On the other hand, high-value or political speech is the central concern of the First Amendment and so is accorded greater protection. It is speech that imparts knowledge or communicates a message.³⁴ Speech without these characteristics is low-value speech. Examples of low-value speech include obscenity, commercial speech and child pornography.³⁵

The distinction between low-value and high-value speech is important because the categories are granted different legal protection. The First Amendment guarantees freedom of expression and prohibits Congress from passing laws abridging freedom of speech.³⁶ However, there are exceptions to the First Amendment and the protection for speech accorded under it differs depending on the type of speech.³⁷ High-value speech is entitled to greater protection, and there is a high threshold before it can be banned. In comparison, low-value speech may be subject to government regulation on the basis of a lower standard and less powerful demonstration of harm.³⁸ In short, low-value speech is only granted limited constitutional protection.³⁹ In the United States, advocates for the regulation of hate speech or pornography must show that the type of speech is not protected under the First Amendment.

Pornography is generally viewed as low-value speech, and thus a central reason why feminist theorists claim its regulation as justified.⁴⁰ Pornography is allegedly low-value because it is non-political—Sunstein describes it as “more akin to a sexual aid than a communicative expression”.⁴¹ Categorising pornography as low-value speech supports the position that the regulation of pornography is consistent with the First Amendment.⁴²

In contrast to pornography, hate speech is normally classified as high-value speech because it often relates to political matters. Mayo Moran notes that in the United States, pornography and racist speech are accorded different constitutional treatment because pornography is viewed as commercial speech, whereas hate speech is classed as

32 Genevieve Lakier “The Invention of Low-Value Speech” (2015) 128 Harv L Rev 2166 at 2168–2169.

33 Cass R Sunstein “Pornography and the First Amendment” [1986] Duke L J 589 at 602–604.

34 At 602–604.

35 Geoffrey R Stone “Content Regulation and the First Amendment” (1983) 25 Wm & Mary L Rev 189 at 194–195.

36 United States Constitution, amend I.

37 Lakier, above n 32, at 2170–2171.

38 Sunstein, above n 33, at 602.

39 Stone, above n 35, at 194.

40 See for example Sunstein, above n 33, at 591.

41 At 606.

42 At 591.

political.⁴³ Courts often view hate speech as being political or cognitive speech—sending a message rather than just being hateful. For example, in *Wall v Fairfax New Zealand Ltd*, the High Court found that the cartoons at issue were contributions to a “public debate about an important issue of public policy”.⁴⁴ This case demonstrates that New Zealand too distinguishes between high and low-value speech: the cartoons were given greater protection because they were deemed to be political expression.⁴⁵

One problem with the distinction between low-value and high-value speech is that political speech is often unclear.⁴⁶ Hate speech does not necessarily always send a political message—sometimes hate speech in a public place is done “for no motivation other than to terrorise”.⁴⁷ It is likewise unclear that pornography is non-political. Feminist theorists have challenged the distinction between public and private spheres,⁴⁸ and have argued that the private is political.⁴⁹ Therefore, feminists would argue that pornography, which is often classified as belonging to the private sphere, *is* in fact political.⁵⁰ Ironically, Sunstein notes that this may weaken the argument for regulation—feminist theorists argue that pornography represents an ideology, and speech that amounts to an ideology may not be considered low-value speech.⁵¹ This contradiction is outside the scope of my article. However, it suffices to say that the distinction between low-value and high-value speech is imprecise.

Hate speech and pornography are not equivalent. Despite this, they have an equivalent function—they both operate to silence. The distinction between high-value and low-value speech does not undermine the argument that hate speech should be regulated because it silences racial minorities. It is not fatal to my argument that pornography and hate speech have differences.

C Comparing sex and race

There may also be some problems with comparing the discrimination directed towards sex and race. One of the difficulties with using the feminist argument in the context of race is that the inequalities faced by sexual and racial groups are not equivalent.

Problems with a sex-to-race analogy arise because women and racial minorities face different experiences of inequality and discrimination. The situation of these groups is fundamentally unique and to have a direct analogy risks ignoring these unique experiences. For example, an intrinsic aspect of racial inequality, particularly in the United States, is segregation. MacKinnon notes “segregation is not the central practice of the

43 Moran, above n 23, at 1430, n 12.

44 *Wall v Fairfax New Zealand Ltd*, above n 31, at [88].

45 At [88].

46 Sunstein, above n 33, at 604.

47 Karl du Fresne “Where should we draw the line between free speech and hate speech?” (23 July 2018) Noted <www.noted.co.nz> per Duncan Webb.

48 Ronnie Cohen and Shannon O’Byrne “‘Can You Hear Me Now...Good!’ Feminism(s), the Public/Private Divide, and *Citizens United v FEC*” (2013) 20 *UCLA Women’s LJ* 39 at 40.

49 Carol Hanisch “The Personal Is Political” in Shulamith Firestone and Anne Koedt (eds) in *Notes From The Second Year: Women’s Liberation* (Radical Feminism, New York, 1970) 76. This essay popularised the slogan “the personal is political”.

50 See, for example, MacKinnon, above n 11, at 195.

51 Sunstein, above n 33, at 607.

inequality of the sexes”.⁵² The history of racial segregation in the United States lies behind the prejudice and discrimination faced by racial minorities today.

Despite this, women and racial minorities face some similar inequalities. These similarities mean that aspects of the silencing argument could apply in the racial hate speech context. Both groups experience stigmatisation, exploitation and denigration based on a characteristic they were born with.⁵³ In addition, both groups face the problem of purported neutrality in society and are at a disadvantage when all groups in society are treated neutrally. The so-called principle of neutrality prevents change when the reality is not neutral.⁵⁴ Both groups have also experienced how differences in society can lead to dominance and an imbalance of power. Women and racial minorities are both systemically disadvantaged from birth,⁵⁵ and without affirmative action, which is often eschewed in favour of neutrality, this disadvantage only increases. Furthermore, there are also those who experience the intersection of race and sex discrimination—in particular, women of colour face the “intersecting patterns of racism and sexism”.⁵⁶ The marginalisation of women of colour means they are at risk of being twice silenced.

Although there are differences in the experiences of racism and sexism, I argue that it is acceptable to compare sex and race in the context of this article. This is because I focus on just one aspect of discrimination—hate speech and pornography—and the experiences of racial minorities and women in relation to this aspect are approximately equivalent. Both hate speech and pornography “enact the abuse”,⁵⁷ objections and responses to hate speech and pornography are routinely trivialised,⁵⁸ and victims are unable to avoid the presence of hate speech and pornography in their everyday physical environments.⁵⁹ The silencing argument is a theoretical argument not limited to describing the experiences of women, but also applicable to the experiences of other marginalised groups, such as racial minorities. If the silencing argument was more closely linked to the unique prejudices that women face, then comparing the experiences of sex and race may have been inappropriate. However, while it is important to point out the potential problems of drawing comparisons, I do not see this as sufficient reason to abandon the application of the silencing argument to racial hate speech.

Having established that the silencing argument *can* be applied to racial hate speech, I shall now show how the argument *does* apply to racist speech.

IV Hate Speech: Silencing Racial Minorities

A *Defining silencing*

Silencing can occur in three main ways: literal silencing, the failure to achieve the intended effects of speech, and illocutionary disablement or the failure to perform an intended

52 Catharine A MacKinnon “Pornography, Civil Rights, and Speech” (1985) 20 Harv CRCL L Rev 1 at 9.

53 At 9.

54 At 9.

55 At 9.

56 Kimberle Crenshaw “Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color” (1991) 43 Stan L Rev 1241 at 1242–1243.

57 MacKinnon, above n 4, at 104.

58 At 104–105.

59 MacKinnon, above n 11, at 183; and Matsuda, above n 2, at 2337.

action. These categories were developed by feminist theorists such as Langton and Caroline West.⁶⁰ Speech can operate to silence people in each of these three ways, albeit to differing extents. I will examine each of these categories, focusing on how speech can silence racial groups in particular.

Defining silencing is important, as many objections to the silencing argument will be based on different interpretations of silencing. For example, Michelman points to a weakness of the silencing argument, in that no literal silencing occurs.⁶¹ It is thus important to use the term silencing with care and precision. This definition of silencing will provide a foundation for the rest of this article.

(1) Literal silencing

The first type of silencing refers to instances where members of a group are literally silenced, in that they utter no words.⁶² This category is closest to the orthodox definition of the term and is the easiest concept to grasp. Literal silencing can be caused in several ways—people may be silent because they are intimidated into silence, or because they believe that nobody will listen, so there is little point in speaking.⁶³ Therefore, speech can silence by creating a social climate that makes groups apprehensive to utter any words.⁶⁴

Hate speech can literally silence by provoking a physical reaction that prevents any speech or reaction. The emotional response to being attacked by hate speech can have a direct physical impact, which precludes speech.⁶⁵ The victim is temporarily disabled by feelings of shock, shame, hurt and fear, which interfere with their capacity to speak.⁶⁶ Abigail Levin terms this a “chilling effect”, which is so powerful that racial minorities cannot rebut racist speech.⁶⁷ This feeling is well known to anyone who has ever been bullied. Often the perfect comeback comes to mind the next day—but at the time, victims are unable to respond. Hate speech creates this feeling on an extreme scale, which leads to the literal silencing of racial groups.

Charles R Lawrence discusses how hate speech can silence racial groups by making speech an inadequate response.⁶⁸ Speech is inadequate to the victim, to the aggressor and to others in society. For the victim, there are no words that can be said to redress the harm caused by being personally attacked by hate speech.⁶⁹ For the aggressor, any spoken response from the victim is likely to be ridiculed and discredited—I expand on this in the next category of silencing. For the observer, if the message of the hate speech corresponds with widely held societal beliefs, any response will likely be ignored or misinterpreted.⁷⁰

The proliferation of hate speech can also literally silence by creating a climate that makes racial minorities less likely to speak. According to feminist theorists, pornography

60 Langton, above n 4, at 314–315; and Caroline West “The Free Speech Argument against Pornography” (2003) 33 *Canadian Journal of Philosophy* 391 at 397–403.

61 Michelman, above n 12, at 296, n 13.

62 Langton, above n 4, at 327.

63 At 327.

64 West, above n 60, at 402.

65 Charles R Lawrence “If He Hollers Let Him Go: Regulating Racist Speech on Campus” [1990] *Duke L J* 431 at 452.

66 At 452.

67 Abigail Levin “Pornography, Hate Speech, and Their Challenge to Dworkin’s Egalitarian Liberalism” (2009) 23 *Public Affairs Quarterly* 357 at 362.

68 Lawrence, above n 65, at 453.

69 At 453.

70 At 453.

can operate to literally silence as its existence has the direct effect of causing women to not speak, protest or express their views—“pornography terrorizes women into silence”.⁷¹ The presence of hate speech acts similarly by creating a climate that can literally prevent minority members from publicly expressing their views.⁷²

The effect of hate speech on a society's climate is difficult to prove, but it is clear that it does have some effect. Hate speech is pervasive throughout society and everyday life. This means that victims of hate speech cannot escape its presence—much like how women cannot avoid the omnipresence of pornography.⁷³ Waldron discusses how symbols like burning crosses or nooses send a terrorising and threatening message to members of racial minorities, making these groups less likely to speak in society.⁷⁴ Richard Delgado similarly notes the silencing effect of hate speech and hateful images such as Confederate flags.⁷⁵ This could be because permanent objects such as flags are a relentless reminder of the harmful message they represent.

The media is one indicator of the climate of the society and the speech that is silenced. In New Zealand media, negative representations of Māori and Pasifika people are prevalent. Negative racial patterns in the media include: constructing Pākehā as the norm, claiming Māori enjoy special treatment, and depicting Māori or Pasifika culture as “primitive”.⁷⁶ These patterns by themselves do not necessarily amount to racial hate speech. However, the depiction of racial minorities in this way contributes to a climate in which minorities feel unable to speak. Negative depictions in media may also lead to further hate speech by consumers, which in turn engenders more media coverage, entrenching the patterns and spreading the message of hate speech.

Another clear example of where hate speech can create a silencing climate is at universities. Universities are renowned for their historic protection of freedom of speech, often valuing freedom of speech over the regulation of hate speech. This is especially harmful as racist speech and conduct is often prevalent at universities. In the United Kingdom, the number of reported racist incidents in universities increased by over 60 per cent from 2016 to 2018.⁷⁷ Punishments for students involved in these incidents were often minor such as warnings, fines or writing a reflective essay.⁷⁸ Students were dissuaded from complaining about racism as all evidence pointed towards there being little repercussion for the offenders.⁷⁹ This is a clear example of the silencing effect.

The presence of hate speech at universities is a current and hotly contested issue in New Zealand. In 2019, white supremacist posters resurfaced on the University of Auckland

71 MacKinnon, above n 11, at 140.

72 Wojciech Sadurski “On ‘Seeing Speech Through an Equality Lens’: A Critique of Egalitarian Arguments for Suppression of Hate Speech and Pornography” (1996) 16 *Oxford J Legal Stud* 713 at 714.

73 MacKinnon, above n 11, at 183.

74 Waldron, above n 5, at 118.

75 Richard Delgado and Jean Stefancic *Understanding Words that Wound* (Westview Press, Colorado, 2004) at 142.

76 Angela Moewaka Barnes and others “Anti-Māori themes in New Zealand journalism—toward alternative practice” (2012) 18(1) *Pacific Journalism Review* 195 at 198, table 1.

77 Eleanor Busby “Racist incidents at UK universities have risen by more than 60 per cent in two years, figures show” *The Independent* (online ed, London, 11 June 2018).

78 Busby, above n 77.

79 Busby, above n 77.

campus, and Vice-Chancellor Stuart McCutcheon refused to remove or condemn them.⁸⁰ This motivated over 400 staff to sign an open letter declaring that racism and white supremacy have no place at the university.⁸¹ As an authoritative body, the University's refusal to remove the racist posters contributed to the narrative that majority groups can spread racist messages without repercussions. Actions (or inactions) such as this shape the climate on campus, and it is easy to see how the groups that these messages were aimed at would feel less welcome on campus or less likely to speak up.

(2) Failing to achieve intended effects

The second type of silencing is where people speak but fail to achieve the effects intended by that speech.⁸² Even if the ideas are communicated, they are not taken seriously and fail to secure agreement.⁸³

According to West, speech (specifically pornography) can silence in this way by producing or reinforcing a social climate in which no one takes the victim's speech seriously—instead ignoring, ridiculing or disbelieving them.⁸⁴ This demeans the victim and reduces their credibility, affecting interpretation of their speech and reducing its value.⁸⁵ West argues that the right to freedom of speech includes the right to not be prevented from distributing meaningful words.⁸⁶ Therefore, freedom of speech is worthless if the speech is made meaningless.

When faced with this response to speech, powerless groups may become convinced that it is not worthwhile to speak as they will not be able to achieve their intended effects. Simultaneously, other groups in society may become convinced that it is not worth listening to this speech.⁸⁷ Thus, this second type of silencing may lead to literal silencing in the future—if speech silences groups by making others ridicule or disbelieve their views, these groups may accordingly speak less to avoid this negative consequence.

There are two possible positions on whether hate speech can cause people to fail in achieving the intended effects of their speech. The first position is that unlike pornography, hate speech does not influence the social climate to such an extent that racial groups are silenced. Feminist theorists argue that pornography convinces its consumers to adopt certain views about women and sex.⁸⁸ Pornography functions as a filter through which both men and women perceive their gender roles and relationships.⁸⁹ This filter effect is likely to influence how people, particularly men, interpret women's speech. Women's speech may be disbelieved, discredited or misinterpreted by consumers and even wider society.

80 Daniel Meech "Vice-Chancellor McCutcheon Calls New Batch of White Supremacist Posters 'Unfortunate', But Says the University Will Not Remove or Condemn Them" *Craccum* (online ed, Auckland, 29 September 2019).

81 "No place for racism: an open letter from University of Auckland staff" (2 October 2019) *The Spinoff* <<https://thespinoff.co.nz>>. See also Kendall Hutt "White supremacists at Auckland Uni: staff sign open letter saying racism has no place there" (2 October 2019) *Stuff* <www.stuff.co.nz>.

82 Langton, above n 4, at 323.

83 West, above n 60, at 402.

84 At 398.

85 Sadurski, above n 72, at 714.

86 West, above n 60, at 408.

87 Sadurski, above n 72, at 717.

88 At 720.

89 Sunstein, above n 33, at 601.

According to some theorists, hate speech does not create a similar filter effect. MacKinnon claims there is no evidence that consumers of hate speech aggress against victims of hate speech, regardless of whether the consumer agrees with the content of the speech.⁹⁰ She claims that pornography works by “circumventing conscious processes”, and hate speech does not.⁹¹ This means that ideas promoted by pornography, such as women being portrayed as less than human, are subconsciously implanted without the consumer being overtly aware that such an idea is being advanced through pornography. In contrast, she claims that hate speech does not work in the same way. If this position is true, hate speech would not silence in the second way because it would not have the requisite effect on the consumers of hate speech in society. Unless consumers were affected, or their mindsets subconsciously subverted, then racial minorities would not have the problem of failing to achieve the intended effects of their speech.

The second, more convincing position is that hate speech does affect the social climate and subconscious thought of those who hear it, at least to the same extent as pornography. An example of hate speech creating a filter effect is where it depicts its victims as less than human. Hate speech often compares racial minorities to universally despised species. In 2018, President Trump described undocumented immigrants as “animals”.⁹² Observers of his speech drew parallels between his comments and “Nazi descriptions of Jews as rats and vermin before the Holocaust”, and Hutus describing “Tutsis as snakes and cockroaches” during the Rwandan genocide.⁹³

These examples show how hate speech can be more overt than pornography in depicting certain groups as less than human. The comparisons to animals are not subtle—they are metaphors rather than similes, declaring that people *are* rats, *are* vermin. This kind of dehumanising speech normalises extreme behaviour and alters what is seen as acceptable speech or behaviour.⁹⁴ This may alter beliefs about minority groups by desensitising them or making other people less likely to speak out in support of the victims of such speech. Therefore, these kinds of hate speech operate to silence in this second way,⁹⁵ as they diminish victim’s abilities to speak and have their speech taken seriously in the public arena.⁹⁶

Wojciech Sadurski’s term “mental intermediation” is another example of hate speech’s effect on social climate.⁹⁷ This intermediation serves to shape our attitudes towards victims in relation to racist speech.⁹⁸ Mari J Matsuda refers to psychosocial and psycholinguistic analysis that exemplifies how racist hate speech has the effect of planting racial inferiority in our minds as one with elements of truth.⁹⁹ Even if we may reject the ideas spread by hate speech, constant exposure to these ideas will bury their way into our subconscious, and “the next time we sit next to one of ‘those people’ the ... message is

90 MacKinnon, above n 4, at 62.

91 At 62.

92 Aliza Luft and Daniel Solomon “How dangerous is it when Trump calls some immigrants ‘animals?’” *The Washington Post* (online ed, Washington DC, 25 May 2018).

93 Luft and Solomon, above n 92.

94 Luft and Solomon, above n 92.

95 Sadurski, above n 72, at 714.

96 Matsuda, above n 2, at 2376.

97 Sadurski, above n 72, at 715–716.

98 Matsuda, above n 2, at 2340.

99 At 2339.

triggered ... interfering with our perception and interaction with the person next to us".¹⁰⁰ Mental intermediation may impact our perception of victims of hate speech, consequently serving to silence these victims.

(3) Illocutionary disablement or failing to perform intended action

The final type of silencing occurs when someone speaks and not only fails to achieve the intended effects of their speech, but actually fails to perform the very action that was intended by their speech.¹⁰¹ Langton terms this "illocutionary disablement".¹⁰² People are prevented from performing an illocutionary act—an act that a speaker performs by speaking the right words in the right contexts.¹⁰³ For example, a same-sex couple saying the words "I do" at a wedding in a place where same-sex marriage is illegal would fail to achieve the action intended by those words.¹⁰⁴ There is a tendency to conflate this type and the second type of silencing into one category, but they do have important, albeit subtle, differences. Illocutionary disablement refers to speech that makes *actions* impossible, whereas the second form of silencing refers to the intended *effects* of speech.

It is difficult to show how hate speech can operate to silence in this third way as there are few clear and concrete examples. Illocutionary disablement is easier to understand in terms of pornography. Pornography silences women's speech in this way by "scrambling" women's speech in sexual contexts".¹⁰⁵ It creates or establishes beliefs in consumers, which cause women's speech and actions in sexual contexts to be systematically misunderstood.¹⁰⁶ For example, the act of refusing to have sex may be made unspeakable because of the systematic belief spread by pornography that when a woman says no, she really means yes. Not only does she not achieve the intended effect of her speech, she also fails to perform the act.¹⁰⁷ Pornography makes actions such as refusal and protest unspeakable for women in sexual contexts.¹⁰⁸ The illocutionary disablement of a victim of rape or sexual assault is an example of how pornography can silence women's speech in this third way.

There is not a similarly obvious example for hate speech—yet it can still silence in this third way by scrambling speech in certain contexts. In fact, hate speech is also relevant to illocutionary disablement in sexual contexts. MacKinnon notes that sexuality is a dynamic in racism, and much racist behaviour and speech is sexual.¹⁰⁹ Hate speech towards women of colour may, together with the effect of pornography, serve to strengthen the conception that women do not mean "no" with the result that women of colour fail to achieve their intended action in sexual contexts.

Furthermore, the mental intermediation that occurs during the second type of silencing may lead to silencing in this third sense. If ideas spread by hate speech are inherently buried in people's minds, then some speech by racial minorities may be rendered impossible. For example, picture an employer who subconsciously believes, as a

100 At 2340 (footnote omitted).

101 Langton, above n 4, at 314–315.

102 At 323.

103 Ishani Maitra "Silencing Speech" (2009) 39 Canadian Journal of Philosophy 309 at 312.

104 Langton, above n 4, at 317.

105 West, above n 60, at 402–403.

106 At 402–403.

107 Langton, above n 4, at 323.

108 At 324.

109 MacKinnon, above n 4, at 63.

result of being consistently exposed to racist speech, that members of a certain race are inherently lazy, dishonest and bad workers. If a person of this race attempts to apply for a job with this employer, any speech uttered by them will fail to count because the outcome is predetermined. Even though the speech is not made legally impossible, such as in the same-sex wedding example, it is also effectively impossible and unspeakable

In conclusion, it is apparent that the silencing argument can apply to hate speech in relation to all three types of silencing. This is not to say that every instance of hate speech serves to silence victims. In some situations, hate speech can lead to a mobilisation of opposition and protest against the speech, or wide societal disapproval of the hateful message. However, the important point is that all instances of hate speech have the *ability* to silence. All three categories of silencing, as established by feminist theorists, are applicable to racial hate speech—even despite the differences between pornography and hate speech.

While the three categories of silencing make up the foundation of the silencing argument, there are other key components to this argument. First, groups can be silenced by being prevented from participating in public life. The second is that powerful groups in society often utilise the power imbalance to silence less powerful groups.

B Silencing: preventing public participation

Another way of looking at how racial minorities can be silenced by hate speech is by reference to Waldron's arguments on how hate speech limits public participation. According to Waldron, principles of democratic citizenship justify prohibiting hate speech.¹¹⁰ He focuses on the harm that hate speech does to dignity, defining dignity as:¹¹¹

... 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.

Hate speech is designed to undermine a person's dignity and undermine their social standing and basic reputation.¹¹²

Waldron further argues that hate speech should be regulated because it prevents its targets from effectively participating in the political life of their society.¹¹³ The right to participate in political and public affairs is a crucial right in a democratic society.¹¹⁴ Participation rights are at the core of eliminating marginalisation and discrimination, and are linked to other rights such as the right to freedom of speech.¹¹⁵ In order to effectively participate in public and political life, a person must have basic social standing and be treated by others as equal. Waldron sees hate speech as a danger to this social status. One of the core aims of hate speech is to compromise the dignity and social standing of its targets—both in their own eyes and in the eyes of others.¹¹⁶ If hate speech is effective in doing so, then victims of hate speech are denied the right to political participation.

110 Waldron, above n 5, at 105 and 108–109.

111 At 105.

112 At 5.

113 At 154–155.

114 United Nations Human Rights: Office of the High Commissioner “OHCHR and equal participation in political and public affairs” <www.ohchr.org>.

115 United Nations Human Rights: Office of the High Commissioner, above n 114.

116 Waldron, above n 5, at 5.

The denial of participation rights can be interpreted as a form of silencing. To be prevented from participating in public and political life equates to being silenced in these spheres. Participation is a form of speech. An attack on a person's dignity or social standing can thus also be seen as an attack on their speech. Hate speech that undermines the reputation of a person or group means that their views may be ignored or ridiculed, which I have already defined as a type of silencing.

The silencing of groups in participatory processes is not just theoretical, but also historically accurate. Iris Marion Young discusses how even if all citizens have the same legal right to participate in decision-making processes, the perspectives of some groups are often silenced.¹¹⁷ Women, racial minorities, working-class and poor people tend to participate less than their counterparts.¹¹⁸ In general, participatory democratic processes tend to silence disadvantaged groups in society.¹¹⁹

This component of the silencing argument also distinguishes the silencing of groups such as women and racial minorities from the silencing of other general groups, such as shy students who are silenced in class discussions by more outgoing children. This is first because of the historic discrimination against the former type of groups that is not present in the latter. The silencing of racial groups and women is intrinsically linked to their constant subjugation and discrimination throughout history. Another distinction is the context in which the groups are prevented from participating. The silencing of women and racial groups prevent them from fully participating in society. Shy students, on the other hand, are merely prevented from participating in classroom discussions, which is a less foundational context.

C *Speech and power*

Power is intrinsically linked with speech, making the existence of power disparities between groups one of the main reasons why speech can operate to silence groups. The interaction between power and speech is clear—the voices of the powerful can silence the voices of the powerless. Power relations also play a part in terms of the various types of silencing. For example, literal silencing may make groups more likely to be intimidated into silence if they also feel powerless. The element of power is crucial—it is difficult to see how any silencing argument would be convincing if two groups in society held similar amounts of power.

Hate speech is effective in silencing racial minorities because of the vastly unequal distribution of power between racial groups in society. Racial minorities face power disparities that affect their ability to speak. Hate speech victims have their speech discounted or silenced because of their lack of power in society,¹²⁰ and hate speech then serves to further entrench this powerlessness.

It is therefore necessary to look at the history of power relations between privileged and subordinated racial groups to understand the effects of hate speech on racial minorities.¹²¹ Lawrence notes that questions of power relationships within which speech

117 Iris Marion Young "Polity and Group Difference: A Critique of the Ideal of Universal Citizenship" (1989) 99 *Ethics* 250 at 258.

118 Jane Mansbridge *Beyond Adversary Democracy* (Basic Books, New York, 1980) as cited in Young, above n 117, at 258.

119 Amy Gutmann *Liberal Equality* (Cambridge University Press, Cambridge, 1980) at 191–202 as cited in Young, above n 117, at 258.

120 Lawrence, above n 65, at 453.

121 Levin, above n 67, at 363.

takes place are crucial considerations when discussing free speech.¹²² The existing power relationships mean that racist speech has a stronger and more damaging impact. Racist speech at an everyday level works with structural racism to “reinforce existing conditions of domination”.¹²³ The relative power between racial groups is one of the primary reasons why the silencing argument holds true.

Disparity in power relations also plays a part in marketplace of ideas, which is often referenced in classic defences of free speech. This concept draws an analogy with the economic marketplace—it is argued that “the test of the truth or acceptance of ideas is dependent on their competition” in an open market, without any censor.¹²⁴ This idea originates from John Stuart Mill’s *On Liberty*, where one of Mill’s primary arguments is that the free competition of ideas, absent of censorship, is the most effective way to separate fact and falsehood.¹²⁵ The marketplace of ideas has evolved to become perhaps the most renowned justification for broad protections of free speech and objections to government regulation of speech.¹²⁶

However, the marketplace of ideas is skewed in favour of the powerful, whose views become established as truth.¹²⁷ Those with the most money and power can buy the most speech.¹²⁸ Those without power cannot buy speech in the marketplace of ideas, and their views are consequently dismissed or devalued. MacKinnon states that “the marketplace of ideas is literal: those with the most money can buy the most speech, and women are poor.”¹²⁹ The theory that all ideas should be heard in society starts from the incorrect assumption that all ideas have an equal starting point. In reality, some ideas are immediately more likely to be believed, spread, and accepted as fact. There is a “maldistribution of private power”,¹³⁰ which means that groups with less power in society are at a disadvantage. Like women, racial minorities have historically held less power in society. This means that for both groups, the entry of their speech into the marketplace of ideas has been at an immediate disadvantage and is less likely to be heard, spread or believed.

One arguable difference between the power imbalances facing groups targeted by hate speech and pornography is that pornography is a well-financed and cohesive industry, and there is no equivalent for racial hate speech. The commercial nature of pornography could mean that there is more of a power imbalance and thus, more of a capacity for the industry to silence women.

However, victims of hate speech may face similar powerful industries. Hate speech is not just isolated and random acts of speech produced by individuals. The most obvious example of an equivalent industry for hate speech is the media. Newspapers, television, radio shows and other media are all effective means to spread messages of, and incite, racial hatred. I have already discussed how media in New Zealand is rife with negative representations of Māori and Pasifika people. One does not have to look far to find other examples of explicitly racist speech in the New Zealand media. In 2003, Paul Holmes used

122 Lawrence, above n 65, at 456.

123 Matsuda, above n 2, at 2335.

124 David Schultz “Marketplace of Ideas” (June 2017) The First Amendment Encyclopedia <www.mtsu.edu>.

125 John Stuart Mill *On Liberty* (Yale University Press, New Haven, 2003) at 88.

126 Schultz, above n 124.

127 MacKinnon, above n 4, at 102.

128 At 102.

129 MacKinnon, above n 11, at 140.

130 Sunstein, above n 33, at 619.

his morning radio show to call the United Nations Secretary-General Kofi Annan a “cheeky darkie” and said “[w]e are not going to be told how to live by a Ghanaian”.¹³¹

In some cases, the private sector or powerful individuals control media industries. Therefore, these media industries risk losing their objectivity by spreading the message of the person or persons in control. For example, in 2003, Rupert Murdoch publicly stated that President George Bush “was ‘acting very morally, very correctly’ by invading Iraq”, following which all 175 of his newspapers published in favour of the war.¹³² The distribution of mass media is an effective way of spreading messages throughout society. There is arguably a power imbalance between the powerful persons who control the media and any hateful messages that are spread, and the victims of any such hateful messages.¹³³

The increased use of social media has also led to an increase in online hate speech, and provided a platform for circulation of extreme and hateful opinions that can reach far more people than previously.¹³⁴ Large media companies such as Facebook and Google can contribute to the power aspect of the silencing argument. Facebook in particular has been criticised for failing to take down or adequately moderate hate speech.¹³⁵ Facebook recently deemed that a graphic post—a graphic image with the caption “the only good Muslim is a f-ing dead one”—did not break their community guidelines after a Facebook user flagged it.¹³⁶ Protests to take down a page called “Jewish Ritual Murder” also failed.¹³⁷ Facebook’s hate speech policy is unclear and lacks transparency—posts such as the above are not deemed hate speech, but statements such as “men are trash” are hate speech.¹³⁸ In 2017, when asked about 49 specific examples of potential hate speech, Facebook admitted that moderators made mistakes on the classification of 22 of these instances.¹³⁹

If these social media giants are not providing sufficient moderation of hate speech, then they are open to exploitation and could become an equivalent industry to the pornography industry. Thus, while there are not exactly equivalent industries producing pornography and hate speech, there are similar industries where parallels can be drawn. The biggest difference is that the pornography industry exists for the purpose of producing pornography, whereas the production and spread of hate speech is not usually the purpose of media and social media sites. Rather, it occurs as a secondary purpose or as side effect. However, the power imbalances that are created by the existence of these

131 Stuart Dye “Outrage at Holmes’ darkie jibe” *The New Zealand Herald* (online ed, Auckland, 27 September 2003).

132 Jeff Sparrow “This is why free speech is an industrial issue” (24 June 2018) *The Guardian* <www.theguardian.com>.

133 Sparrow, above n 132. See also, for example, Charlie Mitchell “Our Truth, Tā Mātou Pono: Over three centuries we’ve failed to represent Māori fairly” (30 November 2020) *Stuff* <www.stuff.co.nz>, which is a recent analysis by Stuff investigating the stereotypes and misrepresentations of Māori perpetuated by newspapers over the years.

134 Fresne, above n 47.

135 See for example, Ariana Tobin, Madeleine Varner and Julia Angwin “Facebook’s Uneven Enforcement of Hate Speech Rules Allows Vile Posts to Stay Up” (28 December 2017) *ProPublica* <www.propublica.org>.

136 Tobin, Varner and Angwin, above n 135 (offensive language censored).

137 Tobin, Varner and Angwin, above n 135. Facebook has taken down the page when Tobin, Varner and Angwin published their article.

138 Mary Meisenzahl “Why Facebook categorizes phrases like ‘men are trash’ as hate speech, according to Mark Zuckerberg” (4 October 2019) *Business Insider Australia* <www.businessinsider.com.au>.

139 Tobin, Varner and Angwin, above n 135.

industries are similar enough that this is not detrimental to the overall application of the silencing argument from pornography to hate speech.

V Objections

The four most likely objections to the preliminary conclusion that hate speech silences racial minorities are as follows. First, that hate speech does not literally silence racial groups. Second, that strengthening free speech in one place strengthens it everywhere. Third, that the state should not engage in content-based regulation and finally, that offensive speech is still entitled to protection. These objections are for the most part based on criticism of the feminist pornography argument.

A *No actual or literal silencing*

The objection most likely to be raised is that racial minorities are not actually silenced by hate speech because speech does not have the power to literally silence people. The presence of hate speech cannot literally prevent people from uttering words.

This objection is prevalent amongst critics of the argument that pornography silences women. Ronald Dworkin calls the silencing argument a “confusion”.¹⁴⁰ Michelman, despite sympathising with the overall feminist argument, describes the silencing as “figurative” and “extended”, and doubts that any real silencing occurs.¹⁴¹ As this objection is not specific to pornography, but relates to the general concept of silencing, it is likely to oppose the argument that hate speech operates to silence racial groups.

Proponents of this objection fail to be persuasive because they rely on the interpretation of silencing as only literal silencing and do not consider the other categories of silencing. Hate speech may not have the ability to literally render somebody mute, but it still has the power to silence. For example, hate speech can create a climate where speech seems impossible, where it will not be interpreted the way that it was intended, or where it results in people being unable to achieve the intended effect or action of their speech. This objection misunderstands the definition of silencing in the silencing argument. Thus, if silencing is understood in a broader sense, this objection is easily refutable.

B *Strengthening free speech*

The next potential objection is that strengthening free speech in one place strengthens it everywhere.¹⁴² Therefore, strengthening the free speech of those producing hate speech also promotes free speech for the victims of hate speech.

This is one of several objections related to free speech, as advocates for free speech will feel the most threatened by the silencing argument. When considering these objections, it is important to keep in mind that free speech is not an absolute right. It is not superior to other democratic rights and values. Under the New Zealand Bill of Rights Act 1990 (NZBORA), all rights are of equal status—the right to freedom of speech is not entrenched, nor is it placed higher than any other right in the Act. This means that a

140 Dworkin, above n 7, at 15.

141 Michelman, above n 12, at 296, n 13.

142 MacKinnon, above n 52, at 4.

possible threat to freedom of speech does not automatically invalidate the silencing argument.

This objection is historically inaccurate—strengthening free speech for some groups has not strengthened free speech everywhere. MacKinnon points out the irrationality of this argument, questioning whether it seems likely that the free speech of Nazis promoted the speech of Jews, or that the speech of the Ku Klux Klan enhanced the speech of black Americans.¹⁴³ It is important to consider which groups in society enjoy the benefits of the protection of free speech. “In the context of our history, we should consider who has most often enjoyed freedom of speech and whose freedom of speech has most often been suppressed or neglected.”¹⁴⁴ In New Zealand, Moana Jackson argues that “[t]he right to free speech ... has too often meant the freedom to hurt, despise and belittle Māori.”¹⁴⁵

Furthermore, the idea that protecting free speech strengthens free speech for all groups in society relies on the problematic principle of neutrality. The principle of neutrality assumes that to treat groups neutrally is to arrive at equality.¹⁴⁶ That is to say, if everyone is subject to the same laws around free speech, then they would be granted equal protection under the law and equal opportunities for speech. However, this argument is flawed because it overlooks the pre-existing disadvantages of certain groups in society.¹⁴⁷ Affirmative action is sometimes necessary to arrive at true equality. Certain racial groups have been historically silenced and having a neutral (and theoretically equal) law does not solve this problem.

It is difficult to see just how free speech of those spreading hateful messages enhances the speech of their targets. Even those wary of regulation (such as Paul Spoonley) question whether free speech is advanced by hateful and extreme views.¹⁴⁸ This objection is therefore unconvincing, as it is historically inaccurate and misrepresents the positioning of groups in society to one another.

C *State being a censor*

Another likely objection is that any government restrictions or regulations on hate speech would be content-based, making the state act as a censor.¹⁴⁹ This is undesirable because of the “slippery slope” argument—that state censorship, once permitted, can quickly lead to censorship of a great deal of speech without justification.¹⁵⁰ This may lead to censorship beyond just the hate speech that such censorship was intended to regulate. Slippery slope arguments are common in discussions about freedom of speech.¹⁵¹

A notorious example of this objection relating to hate speech is the *National Socialist Party of America v Village of Skokie* case in the United States.¹⁵² This case concerned a proposed Nazi march through Skokie—a town in Illinois with a large population of

143 MacKinnon, above n 11, at 209.

144 Fresne, above n 47, per Stacey Morrison.

145 Moana Jackson “No one’s exercise of free speech should make another feel less free” (6 May 2018) E-Tangata <<https://e-tangata.co.nz>>.

146 Cass R Sunstein “Neutrality in Constitutional Law (with Special Reference to Pornography, Abortion, and Surrogacy)” (1992) 92 Colum L Rev 1 at 5.

147 At 5.

148 Fresne, above n 47, per Paul Spoonley.

149 Matsuda, above n 2, at 2351.

150 Frederick Schauer “Slippery Slopes” (1985) 99 Harv L Rev 361 at 361–363.

151 At 363.

152 *National Socialist Party of America v Village of Skokie* 432 US 43 (1977).

Holocaust survivors.¹⁵³ The United States Supreme Court held that all restrictions and limitations that had been placed on the march were unconstitutional.¹⁵⁴ The Court was concerned with legislating against content of First Amendment activity—that is, the right to free speech. The fear was that denying the Nazi marchers free speech protection would create a slippery slope that could result in the denial of free speech protection to those who should theoretically be protected.¹⁵⁵

Jurisprudence and literature have coined this concern the “chilling effect”: where “individuals seeking to engage in activity protected by the First Amendment are deterred from doing so by governmental regulation not specifically directed at that protected activity”.¹⁵⁶ The risk is that the free speech of minority groups, including racial minorities, may be harmed through enacting government regulation intended to protect those very groups. This may include chilling speech such as criticism against majority colonial practices.

The first issue with this objection is that the slippery slope argument is a weak basis for grounding any argument. It is tenuous to argue against something purely on the basis that it might lead to something undesirable. “It is always a doubtful course, to argue against the use or existence of a power, from the possibility of its abuse.”¹⁵⁷ The slippery slope argument is also a fairly rudimentary argument to make as it is difficult to disprove—one cannot disprove a hypothetical future negative consequence.

This objection also assumes that content-based censorship is undesirable, when it is in fact not universally accepted as problematic.¹⁵⁸ Many of the protests against the regulation of pornography and hate speech come from an American standpoint, where such regulation may be deemed unconstitutional.¹⁵⁹ However, countries like New Zealand and the United Kingdom already have some legislation regulating hate speech, and other content-based provisions in the law.¹⁶⁰

It is also arguable that with the current limited regulation of hate speech, the state is already acting as a censor and in effect censoring the speech of silenced racial groups. Censorship can occur not only through regulation but also through governmental inaction. There is notably no uproar over this form of censorship. Protests arise only when those who already have the power of speech feel threatened. Therefore, even if content-based censorship is deemed undesirable, this is arguably already occurring under the status quo.

Finally, the “chilling effect” objection is crucial as a consideration when undertaking any review of hate speech laws, including in New Zealand’s imminent review.¹⁶¹ The discussion and any resulting regulations will need to be carefully balanced to ensure that it is only hate speech that is regulated, and free speech in other areas is maintained. With consultation and advice my view is that this balance can be struck. There are differences between hate speech towards racial minorities and criticism towards majority groups—

153 At 43.

154 At 44.

155 Schauer, above n 150.

156 Frederick Schauer “Fear, Risk and the First Amendment: Unraveling the Chilling Effect” (1978) 58 B U L Rev 685 at 693.

157 *Martin, Heir at law and devisee of Fairfax v Hunter’s Lessee* 14 US 304 (1816) at 344.

158 Dworkin, above n 7, at 13.

159 At 13.

160 See, for example, the Race Relations Act 1965 (UK). New Zealand’s provisions on hate speech are discussed in Part VI of this article.

161 See Part VI of this article for further discussion on the review of New Zealand’s hate speech legislation.

differences in the purpose or intention of the message, and differences in the consequences on both the target of the message and on society in general. Any criteria for hate speech regulations must address these differences.

D *Offensive speech still entitled to protection*

A final objection is that speech classed as offensive or speech we do not like remains entitled to protection. Ronald Dworkin acknowledges that “[p]ornography is often grotesquely offensive”, but argues “we cannot consider that a sufficient reason for banning it without destroying the principle that the speech we hate is as much entitled to protection as any other”.¹⁶² This is a long-standing constitutional principle in the United States. Justice Oliver Wendell Holmes Jr wrote that:¹⁶³

... if there is any principle of the Constitution that more imperatively calls for attachment than any other it is the principle of free thought—not free thought for those who agree with us but freedom for the thought that we hate.

This objection is flawed because it reduces the harm of hate speech and pornography to mere offensiveness. MacKinnon notes such statements “trivializ[e]” the actual harm caused, urging victims to simply “avert their eyes, lock their doors stay home, stay silent”.¹⁶⁴ But the negative effects of exposure to hate speech for victims include experiences of psychological symptoms and emotional distress—from the relatively minor effects of anxiety and nightmares, to self-harm and suicide.¹⁶⁵ Racist speech injures the dignity and self-regard of victims, causing them to doubt their own worth and value.¹⁶⁶ Furthermore, hate speech causes its victims to be silenced, which is a harm in and of itself. I will discuss more about the harm that hate speech causes in Part VI. For now, it is sufficient to say that hate speech should not fall within the category of offensive speech that is still entitled to protection.

The claim that hate speech is merely “offensive” is a claim only ever made by those who are not the targets of hate speech. Waldron makes the point that toleration of hate speech is overwhelmingly easier for bystanders than for the targets of the speech. This is because harm is done to the victims of speech, not “to the white liberals who find the racist invective *distasteful*”.¹⁶⁷ People who are not directly affected by everyday prejudice, discrimination and hateful speech often minimise its effect.¹⁶⁸ For example, Matsuda notes that the responses of target-group members and non-target-group members differ dramatically when it comes to hate speech.¹⁶⁹ Target-group members are likely to respond to an incident with alarm and calls for redress. In comparison, non-target-group members normally disassociate and consider the incidents as “isolated pranks” rather than indicative of a manifest pattern of racial tension.¹⁷⁰

162 Dworkin, above n 7, at 13.

163 *United States v Schwimmer* 279 US 644 (1929) at 654–655.

164 MacKinnon, above n 4, at 105.

165 Matsuda, above n 2, at 2336.

166 Delgado, above n 23, at 135–137.

167 Waldron, above n 5, at 33 (emphasis added).

168 Joe Higham “Invisible violence: Why we need a hate crime law in New Zealand” (16 November 2017) The Spinoff <www.thespinoff.co.nz>.

169 Matsuda, above n 2, at 2327.

170 At 2327.

Even the actual presence of racial hate speech is sometimes only recognised by racial minority groups. Lawrence notes that “[w]e often hear racist speech when our white neighbours are not aware of its presence.”¹⁷¹ To the majority, racist speech or conduct is not visible because it is the status quo.¹⁷² This shows how it is easy for mere observers of hate speech to either ignore racist speech or to class it as merely offensive and therefore still entitled to protection. To the targets of hate speech, however, hate speech is more than merely offensive.

Thus, my preliminary conclusion still stands—that racial hate speech operates to silence its victims. I will now turn to whether the silencing argument is sufficient justification for stronger regulation of hate speech in New Zealand legislation.

VI Justification of Hate Speech Laws

The question of whether hate speech should be regulated is, and has always been, contentious. It involves limitations on freedom of speech—arguably “the most important cornerstone of a liberal democracy”.¹⁷³ There are three main reasons why the silencing argument can provide an answer to this issue. First, silencing in itself is a form of harm caused by hate speech. Secondly, silencing cuts across the right to freedom of speech. This means the objection that hate speech regulation cuts across the right to freedom of expression is untenable. Finally, the silencing effect of hate speech affects not only victims but also impacts how other people in society view those victims.

Before moving onto these main points, I will provide a recap of the relevant law in New Zealand and the recent calls for review of hate speech legislation. This will provide context for where legislative regulation could occur and why there is a current impetus for change.

A *Hate speech in New Zealand*

(1) Legislative provisions

(a) Human Rights Act 1993

Racial hate speech is currently regulated in New Zealand by ss 61 and 131 of the Human Rights Act 1993 (HRA). These sections are the most logical places to make changes. Section 61 treats hate speech as a civil matter whereas s 131 deals with the criminalisation of hate speech. Interestingly, the term hate speech is not explicitly used in either provision. Instead, both provisions uses the term “racial disharmony”.

Section 61(1) of the HRA provides that it is unlawful to publish, distribute or use in any public place, “words which are threatening, abusive or insulting”, and which are:

... 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.

171 Lawrence, above n 65, at 435.

172 At 443.

173 Sam Hurley “Free speech v hate speech: Whose rights do we stand up for?” *The New Zealand Herald* (online ed, Auckland, 3 June 2018) per Muir J.

Complaints of a breach of s 61 are brought to the Human Rights Commission (HRC). The HRC does not have the power to make a ruling but offers mediation,¹⁷⁴ and if complainants are dissatisfied with this, the next avenue is the Human Rights Review Tribunal (HRRT).¹⁷⁵

The number of complaints the HRC receives about racist speech is increasing every year. In 2018, 458 complaints about racism were made to the HRC—the highest number of complaints in five years.¹⁷⁶ These complaints involved instances such as a supervisor using “derogatory names to workers from the Pacific Islands” and overseas workers in Christchurch being banned from bars because of their race.¹⁷⁷ There was also a spike in race-related complaints during the Covid-19 pandemic, with the victims predominantly being people of Chinese and Asian descent.¹⁷⁸ The HRC and HRRT have not been able to cope with the increasing number of complaints and in 2018, the HRRT admitted a backlog of cases of up to two years.¹⁷⁹ This highlights the need for stronger regulation—especially since the number of racist comments and instances may be higher given that not all people who experience racism would make a formal complaint.

The regulation of hate speech under s 61 is also problematic because the required threshold is very high. Some seemingly clear racist speech may not be deemed hate speech under the provision. For example, *Wall v Fairfax New Zealand Ltd* was the first case brought to court in New Zealand involving a complaint under s 61.¹⁸⁰ The case concerned the application of this section in relation to a cartoon that depicted stereotypical and allegedly insulting images of Māori and Pasifika people.¹⁸¹ In this case, the High Court held that the prohibition in s 61 is limited to “relatively egregious examples of expression which inspire enmity, extreme ill-will or are likely to result in the group being despised”.¹⁸² The cartoons in question were held to not meet the requisite threshold of exciting hostility or bringing into contempt Māori and Pasifika people.¹⁸³

Section 131 of the HRA, which deals with the criminalisation of hate speech, has a similarly high threshold. It is a criminal offence for any person to commit the same act(s) as set out in s 61, with the added requisite intention “to excite hostility or ill-will against, or bring into contempt or ridicule, any group of persons”.¹⁸⁴ This provision has been rarely used. There is only one reported prosecution, *King-Ansell v Police*,¹⁸⁵ although King-Ansell was prosecuted under the Race Relations Act 1971 using the predecessor section to s 131.

174 Human Rights Act 1993, s 76.

175 Section 92B.

176 Te, above n 31.

177 Te, above n 31.

178 Eleisha Foon “Spike in racism during pandemic, Human Rights Commission reports” (3 May 2020) Radio New Zealand <www.rnz.co.nz>.

179 Nikki Macdonald “Justice denied – Human Rights Tribunal claim delays balloon to two years” (27 July 2018) Stuff <www.stuff.co.nz>.

180 *Wall v Fairfax New Zealand Ltd*, above n 31, at [3].

181 At 471.

182 At [56].

183 At [86].

184 Human Rights Act 1993, s 131(1).

185 *King-Ansell v Police* [1979] 2 NZLR 531 (CA).

(b) Harmful Digital Communications Act 2015

Online hate speech in New Zealand is governed by the Harmful Digital Communications Act 2015. However, this Act protects individual victims, and does not apply to hate speech against groups of persons. It is a criminal offence under the Act if:¹⁸⁶

- (a) the person posts a digital communication with the intention that it cause harm to a victim; and
- (b) posting the communication would cause harm to an ordinary reasonable person in the position of the victim; and
- (c) posting the communication causes harm to the victim.

(c) Sentencing Act 2002

Hate speech is also relevant at sentencing—but this does not help racial groups who wish to make a complaint or bring a case about hate speech. The Sentencing Act 2002 provides that when sentencing or dealing with an offender, the court must take into account aggravating factors. This includes whether “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”.¹⁸⁷ The extent to which this principle is used is unclear “as the data is not collected”.¹⁸⁸

(d) Broadcasting Act 1989

The Broadcasting Act 1989 places a responsibility on broadcasters to maintain standards that are consistent with “the observance of good taste and decency”, and the principle that controversial issues should be presented from all points of view.¹⁸⁹ The function of the Broadcasting Standards Authority includes to encourage the development of broadcasting codes that safeguard against the portrayal of persons “in a manner that encourages denigration of, or discrimination against, sections of the community on account of ... race”.¹⁹⁰

Broadcasting standards have been developed in relation to denigration and discrimination.¹⁹¹ However, these standards place emphasis on the importance of freedom of expression and require “that a high level of condemnation, often with an element of malice or nastiness” is present to breach the denigration and discrimination standard.¹⁹² Further, these standards are limited in scope as they relate only to film, videos, publications and broadcasting.¹⁹³

186 Harmful Digital Communication Act 2015, s 22(1).

187 Sentencing Act 2002, s 9.

188 Michelle Duff “Criminalising hate speech: New Zealand considers policing hateful expression” (28 April 2019) Stuff <www.stuff.co.nz>.

189 Broadcasting Act 1989, s 4.

190 Section 21(1)(e)(iv).

191 See, for example, Guidelines 6a–6d discussing denigration and discrimination standards in television programmes: *The Free-to-Air Television Code* (Broadcasting Standards Authority, May 2020) at 38–39; and see also, Guidelines 6a–6d discussing denigration and discrimination standards in radio programmes: *The Radio Code* (Broadcasting Standards Authority) at 26–27.

192 See Guideline 6b in both *The Free-to-Air Television Code*, above n 191, at 39; and *The Radio Code*, above n 191, at 27.

193 Human Rights Commission *Kōrero Whakamauāhara: Hate Speech* (December 2019) at 25.

(e) Defamation laws

Finally, while individuals have limited protection against hate speech under general defamation laws, this is unlikely to provide adequate protection for racial groups against hate speech. Protection extends only to individuals and not groups whereas racial hate speech is usually directed to groups. Golriz Ghahraman MP demonstrated the illogical nature of this:¹⁹⁴

[C]alling [an individual] a pedophile for example, that's unlawful, but if you say Muslims or gay men are pedophiles, that will harm them in the same way ... but there's nothing saying that's illegal.

Thus, the primary legislations currently regulating hate speech in New Zealand do not provide sufficient protection for racial groups.

(2) Review of hate speech laws

The need for a review of New Zealand's current hate speech laws has recently become a more urgent priority. The silencing argument could gain traction in the wake of this review. Following the Christchurch terrorist attack in March 2019, the then Minister of Justice Hon Andrew Little MP announced that he would fast-track a review of New Zealand's hate speech law as the current law is "very narrow".¹⁹⁵ The HRA was already in line for review, but was moved forward given the circumstances. Mr Little said that he expected a proposal in relation to hate speech by the end of 2019.¹⁹⁶ In June 2020, Mr Little confirmed Labour Party was still discussing the issue with its coalition partners and that progress on any legislative reform was paused until after the election.¹⁹⁷ At the end of that year, the Royal Commission of Inquiry into the Terrorist Attack on Christchurch Mosques made 44 recommendations, including reforming New Zealand's legal system to adequately deal with hate crime and hate speech.¹⁹⁸ Specifically, the Commission recommended that the government amend legislation to create "hate-motivated offences" in the Summary Offences Act 1981 and the Crimes Act 1961.¹⁹⁹ The Government agreed "in principle" with all 44 recommendations.²⁰⁰ Minister Little was appointed to coordinate the implementation of the recommendations.²⁰¹ In relation to the recommendations on reforming hate speech laws, Prime Minister Jacinda Ardern noted that consultation testing the proposals would be necessary prior to any legislative change.²⁰²

194 Thomas Coughlan "What will happen to our hate speech laws?" (15 April 2019) Newsroom <www.newsroom.co.nz>.

195 "Current hate speech law 'very narrow' - Justice Minister Andrew Little" (3 April 2019) Radio New Zealand <www.rnz.co.nz>.

196 "Little plans fast-track review of hate speech laws" (30 March 2019) Radio New Zealand <www.rnz.co.nz>.

197 Collette Devlin "Hate speech law stalled until after election - no support yet from NZ First" (24 June 2020) Stuff <www.stuff.co.nz>.

198 *Ko tō tātou kāinga tēnei: Report of the Royal Commission of Inquiry into the terrorist attack on Christchurch masjidain on 15 March 2019* (November 2020) at 762-764.

199 At 762.

200 Jacinda Ardern "Prime Minister's comments on Royal Commission of Inquiry into Christchurch Terror Attack" (8 December 2020) <www.beehive.govt.nz>.

201 Ardern, above n 200.

202 Ardern, above n 200.

While the Christchurch attacks gave new impetus to review New Zealand's legislation, calls for review were already widespread. In 2017, the Committee on the Elimination of Racial Discrimination reviewed New Zealand's compliance with the International Convention on the Elimination of All Forms of Racial Discrimination.²⁰³ It stated New Zealand's "existing legislation may be inadequate",²⁰⁴ and recommended for the Government to "review the adequacy of current legislation in addressing and sanctioning racist hate speech".²⁰⁵ In 2017, Human Rights and Race Relations Commissioner Dame Susan Devoy also called for a review of hate speech law.²⁰⁶ Internationally, the Black Lives Matter movement, which gained traction in 2020 following the killing of George Floyd by police officer Derek Chauvin,²⁰⁷ also provides impetus for a renewed focus on hate speech legislation.

The feeling that New Zealand's laws need review is not universally held, and in some instances has faced vehement objection. In response to Devoy's call for review, 27 high-profile New Zealanders wrote an open letter about the threat of such laws to freedom of speech.²⁰⁸ In 2005, a parliamentary select committee launched an inquiry in Auckland about introducing a legislation which would ban New Zealanders from expressing certain views.²⁰⁹ This was met with a strong backlash.²¹⁰ It is interesting, however, to note the framing of the question. If the question was framed in a different sense—such as the possibility of introducing legislation to protect certain New Zealanders from harmful speech—this may have had a more positive reception.

The hate speech law review provides the opportunity to present arguments, such as the silencing argument, as to why greater regulation of hate speech is necessary. The review is likely to focus on the possible expansion of hate speech laws to include speech directed to a group of persons on the basis of religion, such as Muslims, as this was the group that was targeted in the Christchurch attacks. At present, hate speech on the basis of religion is not prohibited under New Zealand's current legislation. It can only be taken into account as a factor in sentencing. While this change is not the focus of my article, the review provides the chance to fully examine New Zealand's hate speech laws. Now is the time for discussion around the harmful effects of hate speech and why greater regulation is justified.

B Silencing argument: justification for regulation?

The silencing argument, as I have interpreted it, provides justification for greater regulation of hate speech. If New Zealand's provisions on hate speech are strengthened

203 Committee on the Elimination of Racial Discrimination *Concluding observations on the combined twenty-first and twenty-second periodic reports of New Zealand* UN Doc CERD/C/NZL/CO/21-22 (22 September 2017); and International Convention on the Elimination of All Forms of Racial Discrimination 660 UNTS 195 (opened for signature 7 March 1966, entered into force 4 January 1969).

204 Committee on the Elimination of Racial Discrimination, above n 203, at [8].

205 At [9].

206 Susan Devoy made several speeches in 2017 on this topic. See for example Jacob McSweeney "Hate starts small": Govt urged to clamp down on hate speech" (27 January 2017) Radio New Zealand <www.rnz.co.nz>.

207 Larry Buchanan, Quoc Trung Bui and Jugal K Patel "Black Lives Matter May Be the Largest Movement in US History" *The New York Times* (online ed, New York, 3 July 2020).

208 Hurley, above n 173.

209 Phil Goff "Goff welcomes hate speech inquiry" (press release, 6 August 2004).

210 Hurley, above n 173.

because of the review, justification is required to withstand objection. Silencing can provide this justification. This is because silencing has three functions that support regulation: it is a harm caused by hate speech, it is a breach of victims' freedom of speech and it has an effect on others in society.

(1) Silencing as a harm

The idea that any regulation requires a harm stems from Mill's harm principle, which states "[t]hat the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others."²¹¹

If the harm principle is interpreted as only involving physical harm, then it is difficult to justify hate speech legislation on the basis that it causes such harm. There are cases that hate speech causes or incites physical violence: for example, evidence suggests that groups exposed to hate speech may be more likely to commit suicide.²¹² However, most hate speech provisions cover more than just speech that causes physical harm, and most harm caused by hate speech is more difficult to quantify.

Theorists are divided about the harm that hate speech causes and have somewhat moved on from Mill's harm principle. For example, Joel Feinberg suggests an "offense principle": where prohibitions are justified because they would be a way of preventing serious offence to others.²¹³ However, as I have discussed in Part V, there are problems with reducing the effect of hate speech to mere offense. Waldron, as we have seen, focuses on the harm that hate speech does to a person's dignity—their "basic entitlement to be regarded as a member of society in good standing".²¹⁴

I contend that silencing in and of itself is a harm caused by hate speech. This is most similar to Waldron's argument that hate speech is an attack on dignity—in Part IV, I discussed how Waldron's argument can be interpreted to fit the silencing argument. Waldron claims that hate speech directed against a group proclaims that members of that "group are, by virtue of their race or some other ascriptive characteristic, not worthy of being treated as members of society in good standing".²¹⁵ This is not unlike the mental intermediation that occurs in the second and third type of silencing, where views and beliefs about racial minorities in society are twisted as a result of hate speech. Being silenced is indisputably a harm. It is something done to people against their will, which limits their opportunities in society. It frustrates people's intentions and actions, deprives people of their right to speak and be heard, and is generally something the majority of people would find injurious in their everyday lives. This harm points towards the need for greater regulation of hate speech.

211 Mill, above n 125, at 80.

212 Brian M Mullan and Joshua M Smyth "Immigrant Suicide Rates as a Function of Ethnophobias: Hate Speech Predicts Death" (2004) 66 *Psychosomatic Medicine* 343 as cited in Dr Naomi Elster "More than Hurt Feelings: The Real Danger of Hate Speech" (5 April 2017) Impakter <www.impakter.com>.

213 Joel Feinberg *The Moral Limits of the Criminal Law* (Oxford University Press, New York, 1985) vol 2 at 1.

214 Waldron, above n 5, at 105.

215 At 106.

(2) Silencing as a breach of freedom of speech

The silencing argument also provides justification for regulation as it can counter the most prevalent objection—that is, any government intervention infringes on citizens' right to freedom of speech. Freedom of speech has historically been viewed as one of, if not, the most important individual right in any democracy.²¹⁶ Anthony Lester and Geoffrey Bindman states that:²¹⁷

Democracy stands ... on the conviction that unpopular ideas should 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.

There are numerous different views about the precise value of freedom of expression and why it should be protected in a democratic society. Mill believes that the existence of freedom of speech would allow us to arrive at truth, as “[w]e can never be sure that the opinion we are endeavouring to stifle is a false [one]”.²¹⁸ Mill's ideas led to the development of the concept of the marketplace of ideas. It is argued that “an individual who seeks knowledge and truth must hear all sides of the question”²¹⁹—all ideas must be allowed to exist in this marketplace without suppression. Joseph Raz focuses on the value of democracy and how the right to freedom of speech both derives from a citizen's ability to participate in the democratic process and is essential for the future survival of democracy.²²⁰

The regulation of hate speech in New Zealand may cut across the right to freedom of expression as codified in s 14 of the NZBORA. Thus, to have a strong argument in favour of greater protections against hate speech, it is necessary to counter the freedom of speech objection and show that any such regulation would not be contrary to NZBORA.

It is here that the silencing argument succeeds in countering the freedom of speech objection in a way that other arguments have previously failed to do. Given that hate speech can be said to silence racial minorities, this means that the non-regulation of hate speech is harmful to the value of freedom of speech. The regulation of hate speech, accordingly, would actually promote freedom of speech by improving the conditions of speech for those who have been silenced by the presence of hate speech.

Thus far, proponents of freedom of speech have focused on the negative impact that state intervention or suppression of speech could have on the freedom of speech, rather than how freedom of speech could be guaranteed by state regulation. MacKinnon states:²²¹

[T]he urgent issue of our freedom of speech is not primarily the avoidance of state intervention as such, but getting affirmative access to speech for those to whom it has been denied.

216 Anthony Lester and Geoffrey Bindman *Race and Law* (Penguin Books, Middlesex, 1972) at 358.

217 At 358.

218 John Stuart Mill *Utilitarianism, Liberty, Representative Government* (JM Dent & Sons, London, 1910) at 79.

219 Thomas I Emerson (1963) “Toward a General Theory of the First Amendment” (1963) 72 Yale LJ 877 at 881.

220 Joseph Raz “Free Expression and Personal Identification” (1991) 11 Oxford J Legal Stud 303 at 308–309.

221 MacKinnon, above n 11, at 195.

MacKinnon's comments flip the typical approach and ask proponents to look at how the status quo could be negatively affecting the freedom of speech of certain groups in society, and whether state intervention could in fact serve to uphold or improve freedom of speech. The regulation of hate speech would serve this purpose.

Other theorists provide similar arguments as to why regulations on speech may also protect freedom of speech. Michelman argues that "privately wrought suppression" of ideas is as real as "governmentally wrought suppression".²²² Opposition to regulation on the basis that such "governmental suppression of political speech" is unconstitutional does not acknowledge that the suppression of ideas by private entities (as in the silencing of women through pornography) is as consequential in inhibiting openness of speech and ideas in society.²²³ Waldron also questions why it is a given that "there needs to be protection only against the constraining laws and never against the racist expression".²²⁴ Thus, some stronger controls surrounding hate speech may be necessary to respect the value of freedom of speech. Regulations on freedom of speech and stronger hate speech laws will enhance the jeopardised freedom of speech of the currently silenced hate speech victims.

This argument could be extended to encourage the government to take positive measures in empowering minority groups to speak alongside strengthening the regulations on restrictions of speech. By actively providing safe platforms or initiatives through which racial groups can share their views, the government will be increasing, not infringing, freedom of speech.

(3) Silencing as having an effect on others

New Zealand's racial disharmony provisions only apply to where speech or actions stir up hostility or contempt amongst other people in society. The focus is on the effect on others and how they consequently view targets of the speech, rather than focusing on the effect of hate speech on victims. Sections 61 and 131 of the HRA relate to speech that "excite hostility against or bring into contempt any [racial] group". The speech by itself, or where there is only an effect on victims, is insufficient to amount to hate speech under the current legislation.

It is arguable that the silencing argument therefore fails because it focuses on the effect that hate speech can have on its victims. However, the second type of silencing involves how hate speech can affect the perception of others on victim's speech through mental intermediation. Hate speech has the effect of creating or reinforcing negative views about minority groups. This affects the interaction of others with members of those groups, even if unconsciously. This may not meet the current existing threshold of inciting hostility or bringing into contempt, but it is clear that the silencing effect does not only affect the victims of hate speech.

The silencing effect is an example of how the threshold under ss 61 and 131 could be lowered while still requiring an effect of hate speech on others. Removing the need for an effect on others entirely could be undesirable, as the law would become more subjective. However, the current threshold has been criticised for being too high as evident through the very few successful prosecutions. A review of the legislation could consider effects such as reducing victims' credibility or devaluing their speech and ideas.

222 Michelman, above n 12, at 303-304.

223 At 303-304.

224 Waldron, above n 5, at 32.

VII Conclusion

Mostly we are silent, and ... when we speak up, nobody listens.

—Andrea Dworkin²²⁵

Andrea Dworkin echoes the sentiments of every person or group who has been silenced by the speech of others. How do we get people to listen? In my view, this is near impossible without legislative reform.

This article has argued that the silencing argument provides a strong justification for greater legislative regulation of hate speech in New Zealand. Hate speech silences its victims. Racial minorities are intimidated and shamed into silence. Their speech is ridiculed, ignored and disbelieved. It fails to count as speech. These effects justify stronger regulation and limitations on freedom of speech.

I have reached the conclusion that hate speech silences racial minorities by drawing on the work of feminist scholars who describe how speech can operate to silence. These theorists particularly focused on how pornography silences women. I have taken the basic tenets of their argument and applied it to racial hate speech.

Although not all elements of the feminist argument perfectly apply to hate speech, the silencing argument is still effective. There are differences in the nature of pornography and hate speech, as well as differences between sex and race and the discrimination faced by these two groups. There are also inconsistencies because of the commercial nature of the pornography industry, but I have drawn an approximate equivalent with the media industry—especially social media. These inconsistencies do not negate the capacity of hate speech to silence racial minorities in all of the ways that pornography silences women.

After dispelling some objections to my above argument, I have concluded that the silencing argument provides a justification for stronger laws regulating hate speech in New Zealand. Hate speech causes harm through silencing and legislation should address this harm. Furthermore, commitment to free speech necessitates affirmative government action to give speech to the silenced. The HRA, which is currently up for review, is the most logical place to make changes to New Zealand's hate speech laws. In balancing the protection of hate speech victims against the right to freedom of speech, I urge participants in the review to consider the silencing effect of hate speech on victims' freedom of speech. Targets of hate speech are silenced by hate speech. The only way to ameliorate this is through government regulation and broader protections. It is time to move from the deliberately *silenced* to the deliberately *protected*.

225 Andrea Dworkin "Silence Means Dissent: Andrea Dworkin on Women & Pornography" *Healthsharing* (online ed, Toronto, June 1984) at 23.