



Free Speech, Hate Speech, and the Problem of (Manufactured) Authority

Sarah Sorial

Abstract

In this paper, I suggest that the concept of incitement as a way of identifying hate speech sometimes locates the harm caused by speech in the wrong sorts of places. Hate speech expressed in the form of “reasoned argument” or academic debate by persons with the relevant authority or expertise potentially causes more harm, though perhaps in less obvious ways. Literature on the concept of authority has demonstrated the way authoritative speakers or speakers with perceived expertise are able to secure uptake for their views. In this paper, I demonstrate how authority and expertise can also be manufactured, enabling speakers to secure uptake in the same sorts of ways as legitimately authoritative or expert speakers. While I am not suggesting legal penalties for speakers who manufacture authority in these ways, I am arguing that we should nevertheless be sensitive to the ways in which this can occur, how it might cause various kinds of harm, and how these harms might be mitigated.

Keywords: free speech, hate speech, incitement, authority, Holocaust denial

Résumé

Dans cet article, je suggère que la notion d’incitation, en tant que critère servant à déterminer si un discours est haineux, attribue de façon parfois erronée le tort causé par le discours. La propagande haineuse, prononcée sous le couvert d’une argumentation logique ou d’un débat académique par des individus possédant une autorité ou des expertises pertinentes, pourrait potentiellement être plus nuisible. La littérature sur le concept d’autorité démontre comment des experts ou des individus dans les positions de pouvoir jouissent d’une certaine facilité à convaincre. Dans cet article, je démontre comment une autorité ou une expertise peut aussi être fabriquée, soit une action permettant d’attribuer autant de pouvoir aux paroles que si les orateurs en question étaient légitimes. Mon intention n’est pas de suggérer que des sanctions juridiques sont nécessaires pour ceux qui falsifient leur autorité. Nous devons toutefois nous pencher sur la question afin d’examiner les conséquences néfastes de telles actions ainsi que de déterminer comment celles-ci peuvent être atténuées.

Mots clés : liberté d’expression, propagande haineuse, incitation, autorité, négation de l’Holocauste

Canadian Journal of Law and Society / Revue Canadienne Droit et Société, 2013, Volume 29, no. 1, pp. 59–75. doi:10.1017/cls.2013.43

1. Introduction

The question of whether hate speech can cause or constitute various harms has been at the centre of the debate about the limits of free speech. Legally, the question is usually resolved by appealing to the category of “incitement.” Employing the language of “incitement” as a way of defining extreme speech captures more overt forms of racism and their role in provoking or inciting violent conduct. It also ensures that speech is more generally protected, because the law does not capture the actual content of the speech act. The argument for speech protection lends *prima facie* support to the claim that only speech inciting violence and hatred should be regulated. Given that overtly inciting speech has a greater propensity to cause harm, it follows that words of this nature ought to attract various criminal and/or civil penalties.¹

There are, however, two significant problems with this approach: first, laws regulating extreme speech on the grounds of *how* it is expressed rather than its content can sometimes be misdirected. Laws are likely to punish speakers who fail to express themselves appropriately due to a lack of education or other necessary skills for speaking in the “right” sort of way. Histrionic or hyperbolic ranting is often characterized as “extreme speech,” even though the effects of such speech may be negligible. By contrast, legal regulation tends to protect those speakers who are able to couch their claims in language that seems acceptable, even though they may cause more harm with their words.²

Second, many racist groups have been able to modify their language in such a way that enables them to avoid capture by legislation. There is emerging evidence that the speech acts of some extreme groups are becoming more sophisticated, polite, and civil.³ As a result, many extremist groups have been able to resist being identified as hate groups or extreme groups, and they are thus protected from prosecution. Moreover, because their racist ideology is increasingly conveyed through civil and respectable language, it has become more acceptable to a wider and more diverse audience.⁴

In this paper, I argue that the concept of incitement as a way of identifying hate speech is inadequate in identifying these more sophisticated types of hate speech, and that this form of racism is often misidentified as either academic debate or political speech. This is not to say that some inciting speech is not harmful. Situations where an inciting speech is given to an angry mob, as described in J. S. Mill’s famous corn dealer example, indicate that harm can sometimes be

¹ On the harms caused and/or constituted by hate speech, see for example MJ Matsuda, “Public Response to Racist Speech: Considering the Victim’s Story” (1989) 87:8 Mich L Rev 2320–81.

² See *Norwood v DPP*, [2004] EWHC 69 (Admin); *Hammond v DPP*, [2004] EWHC 69 (Admin).

³ For example, the extreme right-wing group the Australia First party employs the language of love and care for white Australian identity and culture to justify its policies of zero-net migration and the abolition of multiculturalism. See G Mason, “The Reconstruction of Hate Language” in K Gelber & A Stone, eds, *Hate Speech and Freedom of Speech in Australia* (Sydney: Federation Press, 2007) at 34–58 [Mason]; and AW Jeremy, “Religious Offences” (2003) 7:33 *Ecclesiastical Law Journal* 127–42.

⁴ Mason, *supra* note 3 at 43; Katherine Gelber, *Speaking Back: The Free Speech Versus Hate Speech Debate* (Amsterdam and Philadelphia: John Benjamins Publishing, 2002) at 19; TA Van Dijk, “Discourse and the Denial of Racism” (1992) 3:1 *Discourse and Society* 87–118.

determined with reference to incitement.⁵ But the reverse is also true: a speech act expressed in a “reasonable” way, or one that uses argument as opposed to vitriol, can also cause harm, though perhaps in less obvious ways.

The causal processes responsible for these subtle kinds of harm are complex, and they can be partly attributed to the interplay between the speaker’s identity, the type of language that he or she uses, and how these factors are assessed and received by the targeted audiences. There is a substantial literature demonstrating the ways in which speaker authority is relevant when assessing harmful speech acts.⁶ Put simply, people are able to do harm with their words not only because of *how* the speech is expressed but also because of *who* they are and the social position they occupy. The speech acts of authoritative speakers can enact norms and impose significant obligations on others to do what the speakers say, or audiences may be more easily persuaded by certain views simply on account of the speaker’s identity.⁷ For these reasons, Ishani Maitra and Mary Kate McGowan have argued that speakers with the relevant authority should be held to different standards of legal and moral accountability and may not enjoy free speech protection for their extreme speech acts.⁸

In the first section, I examine the concept of incitement and how it operates in various legislative frameworks. In section two, I demonstrate the ways in which some forms of racism might be classified as either academic debate or political debate by examining two case studies: Holocaust denial and civil racism by extreme right-wing groups. In the final section, I develop an analysis of these cases and the harms they cause with reference to the idea of authority and expertise in general and to manufactured authority in particular.

2. Extreme Speech and Incitement

Extreme speech has been defined as speech that passes beyond the limits of legitimate protest.⁹ It includes speech that advocates violence as a way of achieving political objectives and “hate speech” against persons or groups. The conceptual problem posed by the regulation of extreme speech—whether it is against the state, a group, or an individual—is how to distinguish hatred from ordinary dislike or disagreement.¹⁰ For this reason, an additional element is usually required

⁵ JS Mill, *On Liberty* (London: Longman, Roberts & Green, 1869) at 119.

⁶ For example, much of the literature on the limits of speech draws on JL Austin’s speech act theory to demonstrate the ways in which speakers with the relevant authority are able to do more things with their words. Rae Langton uses these arguments to demonstrate the ways in which pornography is authoritative, while Maitra and McGowan demonstrate how authoritative persons are able to enact norms and impose obligations on others to do what they say. See JL Austin, *How to Do Things with Words* (London: Oxford University Press, 1962); I Maitra & MK McGowan, “The Limits of Free Speech: Pornography and the Question of Coverage” (2007) 13:1 *Legal Theory* 41–68 [Maitra & McGowan]; R Langton, “Speech Acts and Unspeakable Acts” (1993) 22:4 *Philosophy and Public Affairs* 293–330.

⁷ MK McGowan, “Conversational Exercitives: Something Else We Do with Our Words” (2004) 27:1 *Linguistics and Philosophy* 93–111; MK McGowan, “Oppressive Speech” (2009) 87:3 *Australasian Journal of Philosophy* 389–407.

⁸ Maitra & McGowan, *supra* note 6.

⁹ J Weinstein & I Hare, “General Introduction: Free Speech, Democracy, and the Suppression of Extreme Speech Past and Present” in I Hare & J Weinstein, eds, *Extreme Speech and Democracy* (New York: Oxford University Press, 2009).

¹⁰ Robert Post, “Hate Speech” in I Hare & J Weinstein, eds, *Extreme Speech and Democracy* (New York: Oxford University Press, 2009) 127.

in order to justify prohibition or regulation,¹¹ and this additional element is typically identified using the category of incitement.

For example, Australia has various laws regulating forms of “extreme speech,” including anti-vilification laws at both the federal and state/territory levels, and laws regulating the “urging of violence”¹² against both the state and groups within the Australian community. With the exception of federal¹³ and Tasmanian¹⁴ anti-vilification laws, all other jurisdictions have both criminal and civil laws. Criminal provisions usually require a public act that incites or is intended to incite “hatred, serious contempt or severe ridicule of a person” in a way that threatens physical harm to persons or their property.¹⁵ Civil provisions require a lower threshold but, typically, the wording of the civil offence characterizes extreme speech in terms similar to those that describe the criminal offence; that is, in terms of speech that “incites” violence.

Similarly, sections 319(1) and (2) of the *Canadian Criminal Code* prohibit the public incitement of hatred: “Every one who, by communicating statements in any public place, incites hatred against any identifiable group where such incitement is likely to lead to a breach of the peace is guilty of an indictable offence and is liable to imprisonment.”¹⁶ Section 13 of the *Canadian Human Rights Act* makes it an offence to “communicate telephonically” in a way that is “likely to expose a person or persons to hatred or contempt by reason of the fact that person or those persons are identifiable on the basis of a prohibited ground of discrimination.”¹⁷ The common theme to these different regulatory frameworks is that they are intended to capture deliberately inflammatory speech or more obvious instances of hate speech. The legislation thus conceives of extreme speech in very narrow terms, and therefore tends to capture more obvious or blatant forms of hate speech.

For example, in Canada, the very few section 13 cases that have gone to the Canadian Human Rights Tribunal (CHRT), and those where the Tribunal has found a breach of the section, have almost always involved expression that, according to Richard Moon, is “so extreme and hateful that it may be seen as advocating or justifying violence against the members of an identifiable group.”¹⁸ Only a few prosecutions have been brought under section 319, and there have been even fewer convictions. Between 1994 and 2004, there were 93 prosecutions under section 319; thirty-two convictions were entered, of which 27 resulted in prison sentences and 5 in conditional sentences.¹⁹ In Australia, 618 complaints were lodged to the Human Rights and Equal Opportunity Commission (HREOC) under the *Racial Discrimination Act 1974* (Cth) in 2001. Of these, 25 resulted in HREOC public

¹¹ *Ibid.*

¹² *Criminal Code Act 1995* (Cth) Division 80.

¹³ *Racial Discrimination Act 1995* (Cth) ss 18B–18F.

¹⁴ *Anti-Discrimination Act 1998* (Tas) s 19.

¹⁵ *Anti-Discrimination Act 1977* (NSW) s 20D; *Anti-Discrimination Act 1991* (Qld) s 131A; *Racial Vilification Act 1996* (SA) s 4; *Racial and Religious Tolerance Act 2001* (Vic) ss 24–25.

¹⁶ *Criminal Code*, RSC 1985, c C-46, s 319 (1) (2).

¹⁷ *Canadian Human Rights Act*, s 13(1).

¹⁸ R. Moon, *Report to the Canadian Human Rights Commission Concerning Section 13 of the Canadian Human Rights Act and the Regulation of Hate Speech on the Internet* (Canadian Human Rights Commission, October 2008) at 2 [Moon].

¹⁹ *Ibid.* at 15.

inquiries or federal court cases.²⁰ The Anti-Discrimination Board of New South Wales (NSW) (ADB) received 1,100 complaints under the *Anti-Discrimination Act 1977* (NSW) (ADA) in each of the previous three financial years (2006–07, 2007–08, and 2008–09); however, only 2% to 3% in each year were complaints of vilification.²¹ In its inquiry into the effectiveness of racial vilification laws in NSW and, especially, of section 20D of the *Anti-Discrimination Act 1977*, which prohibits serious racial vilification, it was noted by the NSW premier that there have been no successful prosecutions under that provision to date.²²

It would seem that the “form” of the speech, or the way it is expressed, determines whether a prosecution will be brought against the speaker. The “content” of the speech itself does not seem to be as relevant. This form/content distinction is apparent in a number of significant cases, the most recent being the cases of *Whatcott v Saskatchewan (Human Rights Tribunal)*,²³ *Whatcott v Saskatchewan (Human Rights Tribunal) at the Court of Appeal*,²⁴ and finally, by the Canadian Supreme Court in *Saskatchewan (Human Rights Commission) v Whatcott*.²⁵ In 2001 and 2002, Whatcott distributed four flyers expressing the view that homosexuals should be kept out of Saskatchewan public schools. Four people who received these flyers in their homes filed complaints with the Canadian Human Rights Commission, alleging that the material promoted hatred against individuals on the basis of their sexual orientation, thus violating section 14 of the *Saskatchewan Human Rights Code*.²⁶ The Tribunal decided that Whatcott did violate section 14,

²⁰ L McNamara, *Regulating Racism: Racial Vilification Laws in Australia* (Sydney Institute of Criminology, 2002).

²¹ *Roundtable on Hate Crime and Vilification Law: Directions and Developments*, Law School, University of Sydney, 28 August 2009.

²² Parliament of New South Wales, “Racial Vilification law in NSW,” last accessed 9 August 2013, <http://www.parliament.nsw.gov.au/Prod/parlament/committee.nsf/0/D15ADEF09EE12ACB CA257AD70011F5B1>.

²³ Saskatchewan Court of Queen’s Bench, 2007 SKOB 450.

²⁴ Saskatchewan Court of Appeal, 2010 SKCA 26.

²⁵ *Saskatchewan (Human Rights Commission) v Whatcott*, 2013 SCC 11. See also *Mugesera v Canada*, where the Canadian Supreme Court described the elements of the section 319(2) offence of wilfully promoting hatred in very narrow terms. The Court defined “promote” as actively supporting or instigating and not simply encouraging. The term “hatred” refers to an “emotion of an intense and extreme nature that is clearly associated with vilification and detestation ... only the most intense forms of dislike fall within the ambit of this offence” (at 104). However, proof is not required that the communication caused actual hatred. More generally, the law’s purpose is to prevent the risk of serious harm caused by hate propaganda; in determining these questions, the court must take into account the audience and the social and historical context of the speech (Moon, *supra* note 18 at 14). Despite these considerations, courts typically interpret the aforementioned legislative provisions in terms of a form/content distinction.

²⁶ Section 14 of the *Saskatchewan Human Rights Code* states:

14(1) No person shall publish or display, or cause or permit to be published or displayed, on any lands or premises or in a newspaper, through a television or radio broadcasting station or any other broadcasting device, or in any printed matter or publication or by means of any other medium that the person owns, controls, distributes or sells, any representation, including any notice, sign, symbol, emblem, article, statement or other representation:

- (a) tending or likely to tend to deprive, abridge or otherwise restrict the enjoyment by any person or class of persons, on the basis of a prohibited ground, of any right to which that person or class of persons is entitled under law; or
- (b) that exposes or tends to expose to hatred, ridicules, belittles or otherwise affronts the dignity of any person or class of persons on the basis of a prohibited ground.

and that the section was a reasonable restriction on Whatcott's rights to freedom of expression and religion.

On appeal to the Saskatchewan Court of Queen's Bench, Kovacj J found that section 14 of the *Code* must be interpreted in accordance with the standard of hatred and contempt as outlined in *Taylor*,²⁷ so as to prohibit only "communication that involves extreme feelings and strong emotions of detestation, calumny and vilification."²⁸ He concluded that the flyers did violate section 14 of the *Code* in the ways described in *Taylor* because of the claims made about homosexuals. In the Saskatchewan Court of Appeal, Smith and Hunter JJA found Kovacj J had erred in reaching his decision because he focussed on specific phrases from the flyers rather than on examining the content and context of the flyers as a whole.

Hunter JA argued that section 14(1) of the *Code* must be interpreted and applied so as to only prohibit communications involving extreme feelings and strong emotions of detestation, calumny, and vilification. She held that language used to debate the morality of an individual's behaviour must be protected to a very high degree, and that the flyers were part of an ongoing public debate about teaching homosexuality in public schools. Similarly, Smith JA argued that questions of sexual morality are linked to both public policy and individual autonomy and so are at the heart of protected speech. She found: "[W]here, on an objective interpretation, the impugned expression is essentially directed to disapprobation of same-sex sexual conduct in a context of comment on issues of public policy or sexual morality, its limitation is not justifiable in a free and democratic society."²⁹ The Canadian Supreme Court reversed this decision. The Court reiterated the definition of hatred as previously set out in *Canada (Human Rights Commission) v Taylor* when interpreting hate speech provisions. In determining whether a speech act is hateful, the courts must first consider whether a reasonable person, "aware of the content and circumstances," would interpret the expression as one that would expose a protected group to hatred. Second, the courts must interpret the legislative term "hatred" or "hatred and contempt" narrowly, to include only those extreme manifestations of the emotion described by the words "detestation" and "vilification." This would exclude expression that, while repugnant and offensive, does not incite a level of abhorrence and rejection that risks causing discrimination or other harmful effects. Third, courts must consider the effect of the expression at issue, irrespective of the speaker's intention.

The Court rejected Whatcott's defense that the flyers were related to the discovery of truth and sexual politics and were therefore protected expression. It found that the purpose of hate speech legislation is to restrict the use of representations likely to expose protected groups to hatred and its effects. The expression captured by hate speech laws is typically of an extreme nature. Interestingly, the Court did not accept Whatcott's argument that this was merely

²⁷ *Canada (Human Rights Commission) v Taylor*, [1990] 3 SCR 892.

²⁸ *Whatcott v Saskatchewan (Human Rights Tribunal)*, 2007 SKQB 450 at 21.

²⁹ *Whatcott v Saskatchewan (Human Rights Tribunal)*, 2010 SKCA 26 at 138.

the expression of a political viewpoint and should thus be protected speech. It held that:

Framing that speech as arising in a “moral” context or “within a public policy debate” does not cleanse it of its harmful effect. Indeed, if one understands an effect of hate speech as curtailing the ability of the affected group to participate in the debate, relaxing the standard in the context of political debate is arguably more rather than less damaging to freedom of expression. As argued by some interveners, history demonstrates that some of the most damaging hate rhetoric can be characterised as “moral,” “political” or “public policy” discourse.³⁰

This is an important point, as much of this speech has been characterized as “political debate,” most recently in the US Supreme Court decision in *Snyder v Phelps*.³¹ This, I suggest, is a mischaracterization or misidentification of hate speech as political debate.

Nevertheless, despite its findings against Whatcott, the Court did maintain a form/content distinction when discussing the question of whether this speech could be categorized as protected political communication. It held that while the “polemicist may still participate on controversial topics that may be characterised as ‘moral’ or ‘political’ ... words matter. In the context of this case, Mr. Whatcott can express disapproval or homosexual conduct and advocate that it should not be discussed in public schools or at university conferences. Section 14(1)(b) only prohibits his use of hate inspiring representations against homosexuals in the course of expressing those views.”³² This was based, in part, on the reasoning in *Kempling v College of Teachers (British Columbia)*, where Lowry JA acknowledged that while Kempling’s published writings included a legitimate political element or a “reasoned discourse, espousing his views as to detrimental aspects of homosexual relationships,” and that “he was, in his more restrained writings, engaged in a rational debate of political and social issues,” his writings sometimes “clearly crossed the line of reasoned debate into discriminatory rhetoric.”³³ As such, his writings, on the whole, were not deserving of a high level of constitutional protection.

The apparent tension in all these cases is that, on the one hand, there is an acknowledgement that hate speech can be interpreted broadly to include speech that does not overtly promote hatred and violence, and that the harms of hate speech can be as subtle and diffuse as the forms it takes. On the other hand, there is a general reluctance to legally regulate speech that may be harmful but is not expressed in an extreme way. There are good reasons for this. As argued in the *Keegstra* dissenting judgment, while it was easy to argue in the *Keegstra*³⁴ case that the statements made by the accused contributed nothing to democratic society, in other cases, it may not be so easy distinguish between speech that is socially valuable and speech that is not. Thus, we should err on the side of caution and permit all speech, irrespective of content.

³⁰ *Ibid.* at 116.

³¹ *Snyder v Phelps*, 562 U.S.__(2011).

³² *Whatcott v Saskatchewan (Human Rights Tribunal)*, 2010 SKCA 26 at 119.

³³ *Kempling v College of Teachers (British Columbia)*, [2005] BCJ No 1288 at 76.

³⁴ *R v Keegstra*, [1990] 3 SCR 697.

While there are good legal reasons for maintaining the form/content distinction, there are good social and political reasons for being vigilant about the ways in which speakers of hate speech manipulate this distinction, and for being sensitive to the kinds of harms this causes. As Moon has argued, the use of censorship by governments should be confined to the narrow category of extreme speech, and any attempts to regulate beyond this to capture more commonplace or even civil forms of racism would be an extraordinary and unjustifiable use of state power. He suggests that because less extreme forms of discriminatory expression are so pervasive, it is difficult to establish clear and effective rules of identification and exclusion. But just because they are so pervasive, it is critical that they be identified and addressed in other non-legal ways.³⁵ My aim in this paper is to draw attention to one form that racism can take, a form of what I refer to as “manufactured” authority or expertise, which is commonly mistaken for a type of academic debate or reasoned political argument.

3. Racism as Academic Debate? The Case of Holocaust Denial

In this section, I examine how the form/content distinction emerges in the cases of *R v Zundel* and *Jones v Toben*, and the ways in which speakers can manipulate this distinction. Holocaust denial is an interesting case study for several reasons: first, the legislation in both Australian and Canadian jurisdictions does not prohibit Holocaust denial itself, but only prohibits those statements that are considered injurious to the targeted groups or victims. This means that the legislation may not capture forms of Holocaust denial that appear in the guise of “reasoned” argument, or which appear to be consistent with prevailing norms of civility. Second, deniers typically claim their work has academic merit as a way of evading capture by the legislation. Third, the harm caused by Holocaust denial is not simply that it offends victim groups. The fact that Holocaust deniers are able to manufacture legitimacy for their views by claiming them to be part of a historical debate means that others might come to accept those views because they consider them credible.

The case of Ernst Zundel has a long and protracted history in the Canadian courts. Zundel was born in Germany and arrived in Canada in the 1950s, where he quickly developed close associations with fascist groups in Canada and elsewhere.³⁶ From 1996 to 2002, the Canadian Human Rights Commission (CHRC) pursued an investigation into Zundel’s Internet activities. During this time, Zundel and his supporters had been exploiting the Internet as a way of disseminating neo-Nazi ideas. The CHRC found against Zundel under section 13(1) of the *Canadian Human Rights Act*, which prohibits telephonic communications of “any matter that is likely to expose a person or persons to hatred or contempt by reason of the fact the person or those persons are identifiable on the basis of a prohibited ground of discrimination.”³⁷ The Tribunal heard a series of technical

³⁵ Moon, *supra* note 18 at 1.

³⁶ See D Fraser, “‘On the Internet, Nobody Knows You’re a Nazi’: Some Comparative Legal Aspects of Holocaust Denial on the WWW” in I Hare & J Weinstein, eds, *Extreme Speech and Democracy* (New York: Oxford University Press, 2009) [Fraser].

³⁷ *Canadian Human Rights Act*, s 13(2).

and legal arguments, including the defense of truth. A full examination of these issues is outside the scope of this paper, so this section will focus on the expert testimony about the language and codes of Holocaust denial.

As part of its investigation into Zundel's activities, the CHRC called on the expert evidence of Professors Gary Prideaux and Frederick Schweitzer. Prideaux is an expert in the field of discourse analysis. Discourse analysis is a way of interpreting any given text or meaning through the use of established linguistic principles, and specific strategies used to disguise the meaning of otherwise contentious content. An understanding of the rhetorical strategies and these general principles enables an interpretation of the text and a determination of the likely impact of the communication.³⁸ Some of the techniques Prideaux identified in Holocaust denial literature included generalization and the use of scare quotes to give an additional layer of meaning beyond that which is obvious; specific vocabulary chosen to reflect the author's view of a particular group or event; use of repetition, which serves to enhance the credibility of the author or to persuade the audience of the veracity of a particular assertion; and the use of coding and metaphor to establish a series of negative associations. In the community of Holocaust deniers, readers of the material share the same linguistic code and are able to decipher that "Zionist" and "Marxist" are code for "Jew," and that "Jew" signifies "bad."³⁹

The Tribunal did read and interpret the material on Zundel's website in light of the expert evidence. It found that the messages conveyed in the documents carried very specific assertions about the character and behaviour of Jews, none of which were good. It found: "Jews are vilified in the most rabid and extreme manner ... these messages create an environment in which it is likely that Jews will be exposed to extreme emotions of detestation and vilification. Based on our view that the Zundel site materials characterize Jews as 'liars, cheats, criminals and thugs' ... we regard it as highly likely that readers of these materials will, at a minimum, hold Jews in very low regard, viewing them either with contempt, scorn and disdain, or hatred, loathing and revulsion."⁴⁰ But the Tribunal also found that, had the material been expressed in a less vitriolic way, it might have decided differently.

The Tribunal also heard evidence that the material on the Zundel site was the healthy expression of one perspective in an ongoing historical debate. Mark Weber, the director of the Institute for Historical Review, explained that revisionists define themselves and the field of Holocaust revisionism as a critique of conventional or official history, and that they play an important role in historical discourse: "[T]heir writing and research should be seen as part of a larger debate, and is to be credited with generating a mainstream historical response. In his view, revisionism is similar to any intellectual exchange and is merely at one end of a continuum of historical perspective."⁴¹ Interestingly, the Tribunal seemed to accept the idea

³⁸ *Citron and Toronto Mayor's Committee v Zundel*, 2002 CanLII 23557 (CHRT) at 122.

³⁹ As Fraser puts it: "[T]he signifying sequence is not likely to subject Jews to hatred and contempt since Jews are already subject to hatred and contempt by those who deploy and invoke it. The hermeneutic circle of Holocaust deniers is a closed one." Fraser, *supra* note 36 at 531.

⁴⁰ *Citron and Toronto Mayor's Committee v Zundel*, *supra* note 38 at 140.

⁴¹ *Ibid.* at 150.

that this is but one version of history, but it rejected the idea that the material produced by Zundel is part of such a legitimate historical debate. It noted:

In any event, even if we accept that here can be legitimate debate on this topic, we have focused on the manner in which the Respondent has expressed his views and not the mere fact that he chooses to engage in this debate. Our conclusion is based on the way in which these doubts are expressed, *and not on the fact that challenges are raised regarding the historical accuracy of these events ...* If this truly were a neutrally worded “academic” debate, our analysis might be quite different.⁴²

The issue in contention here is not the debate about Holocaust denial, but the manner in which that debate is carried out. The Tribunal seems to be suggesting here that had Zundel expressed his views in more tempered or “reasonable” language, he may have avoided violating section 13(1) of the *Act*. Consequently, the decision seems to permit Holocaust denial as a matter of legitimate historical debate, provided that debate is conducted in a civil and academic way.

The issue of how the ideas are expressed rather than their content also arose in the Australian case of *Jones v Toben*.⁴³ Fredrick Toben is also an expatriate German resident and the founder of the Adelaide Institute, the focal point of Holocaust denial in Australia. Like those against Zundel, the legal proceedings against Toben have been complex and protracted. Toben came before the federal court after the Human Rights and Equal Opportunity Commission (HREOC) found that he had violated section 18C of the *Racial Discrimination Act* by posting Holocaust denial material on his website. Like the Zundel case, the Toben case addressed the way the Internet was used to disseminate Holocaust denial material, the way in which that material was represented, and the question of truth with respect to whether the Holocaust did in fact occur.

The federal court seemed to take seriously recent work on the nature of speech and harm and, especially, the idea that harms caused by speech need not be in the form of physical violence against persons or their property. The court found that harms could be diffuse and could take the form of silencing victims of hate speech or undermining democratic values and institutions. But it also seemed that what was at issue was not the content of Toben’s speech, but how Toben’s meaning was expressed. Carr J argued:

In the context of knowing that Australian Jewish people would be offended by the challenge which the appellant sought to make, *a reasonable person acting in good faith would have made every effort to express the challenge and his views with as much restraint as was consistent with the communication of those views.* In my opinion, the Document shows that the appellant made no such effort. On the contrary, the terms of the Document are, in my view, *deliberately provocative and inflammatory*.⁴⁴

This implies that had Toben “made every effort” to express his views in a manner other than the deliberately provocative and inflammatory way in which he did, he may have been able to avoid violating the *Racial Discrimination Act*.

⁴² *Ibid.* at 153–54 (emphasis added).

⁴³ *Jones v Toben* (Corrigendum dated 20 April 2009), [2009] FCA 354 (16 April 2009).

⁴⁴ *Ibid.* at 22 (emphasis added).

This is a serious issue, because as Fraser's analysis points out, rarely do the authors of Holocaust denial literature express themselves overtly in statements like "kill the Jews." Many Holocaust deniers, including David Irving, are rarely so blatant and obvious in their discourse. Holocaust deniers typically do not engage in inflammatory or overtly provocative speech; nor do they identify the targets of their speech in a direct way. More often than not, their views are articulated in a vague or coded way. As Fraser notes, simply by replacing references to "Jews" with terms like "our traditional enemies," a denier can effectively avoid regulation. The problem is that if racist views are expressed using civil language, they might easily be mistaken or misidentified as a form of academic debate or as a serious contribution to political discourse, rather than as a more sophisticated type of racism. The danger here is not simply that such speech will offend targeted groups, but that it might make those views seem acceptable and credible to a wider or less discerning audience.

4. Racism and Civility

Extreme right-wing groups in Australia are employing similar techniques in order to evade capture by legislation and appeal to a wider and more diverse audience. For example, the Nationalist Alternative represents itself as an organization committed to the "welfare and needs of the Australian people." Its aim is to "reaffirm Australian cultural and national identity and restore the sovereignty and independence of the Australian nation."⁴⁵ The organization uses the language of autonomy, respect, and nationhood to mask what is, essentially, an anti-immigration and racist agenda: "Our vision of Australia is one of an organic nation, founded upon Western/European ideals, and created by its descendants primarily the Anglo-Saxon-Celtic ethnicity as well as fellow Europeans from northern, central, southern and eastern Europe."⁴⁶ The group's website contains a number of articles and public comment forums that appear as reasoned, reasonable, and civil arguments in defense of a white Australia.

For example, an article entitled "'White Flight' from schools" comments on the alleged problems created by the overrepresentation of Asian and Indian students at selective schools in Australia, and how Asian ways of learning are adversely changing the Australian education system and, by implication, Australian society.⁴⁷ The article references academic papers on different learning styles and manipulates the research to support the author's agenda. The author also extensively references a variety of anecdotal evidence as found in online discussion forums, to give the impression that the research is further supported by white parents' experiences of the selective-school education system. The Australia First party uses similar techniques to lend credibility or legitimacy to its views, including hosting the annual Sydney Forum, a conference-like event that gathers various academics and published authors. Topics covered include free speech, immigration, people smuggling,

⁴⁵ Nationalist Alternative, "About Us," last accessed 9 August 2013, <http://www.natalt.org/about/>.

⁴⁶ *Ibid.*

⁴⁷ M Kennedy, "'White Flight' from schools," Nationalist Alternative, October 28, 2011, <http://www.natalt.org/2011/10/28/%E2%80%9Cwhite-flight%E2%80%9D-from-schools/>.

and Zionism.⁴⁸ The purpose of having “expert” views represented at these forums is to make the organization look credible, and to make the arguments it espouses look like legitimate political debate. I am not suggesting that debate about these issues is not legitimate; my concern is with the way such views are manipulated to further what is, essentially, a racist agenda, and how some undiscerning audiences might receive such views.

5. Manufacturing Authority/Legitimacy and the Question of Harm

In *R v Keegstra*,⁴⁹ Dickson CJ referred to a quote from the report of the Cohen Committee, which had found that “individuals can be persuaded to believe almost anything if the information or ideas are communicated using the right technique and in the proper circumstances ... It is thus not inconceivable that the active dissemination of hate propaganda can attract individuals to its cause.”⁵⁰ For these reasons, Dickson CJ concluded that the state is justified in restricting the extreme expression of people like Keegstra, because such expression might cause others to hate the members of the targeted group and to act towards those members in a violent or discriminatory way. It might also cause the targeted group members to internalize hateful views, thereby damaging their self-esteem. This causal argument between expression and the spread of hatred is based on a mistrust of the role of rational agency in certain circumstances. Dickson CJ noted that we should not “overplay the view that rationality will overcome all falsehoods in the unregulated marketplace of ideas.”⁵¹

Recent empirical research about how people form beliefs and weigh reasons for acting has confirmed this argument. For example, studies have shown that people tend to accept propositions they understand, even when they are explicitly told that the propositions are false. As Daniel Gilbert argues, due to the way in which we form beliefs, we have a tendency to automatically accept propositions we can easily understand; we are also poor at ignoring, forgetting, rejecting, or otherwise failing to believe what we have comprehended, and we tend to believe information to be true even when we cannot determine its truth value by assessment. Rejection requires effort, and this effort can be impaired in any number of ways, including sleep deprivation, torture, attention deficits, intellectual capacity, and education.⁵² The empirical evidence thus indicates that we are especially prone to accepting as true the things we see and hear, and that we do not always rationally engage with what we hear. This means that being exposed to more or “better” arguments will not necessarily enable us to come to the right decision or lead us to change our erroneous views; we may not even be any more inclined to understand different viewpoints.⁵³

⁴⁸ Australia First Party, “The 2006 Sydney Forum: 26–27 August 2006,” last accessed 9 August 2013, <http://www.australiafirst.net/2006/news200607sydneyforum.htm>.

⁴⁹ *R v Keegstra*, [1990] 3 SCR 697.

⁵⁰ M Cohen, *Report to the Minister of Justice of the Special Committee on Hate Propaganda in Canada* (Ottawa: Queen’s Printer, 1966), cited in Moon, *supra* note 18 at 22.

⁵¹ *R v Keegstra*, [1990] 3 SCR 697, per Dickson CJ cited in Moon, *supra* note 18 at 22–23.

⁵² D Gilbert, “How Mental Systems Believe” (1991) 46:2 *American Psychologist* 107–19.

⁵³ This is further substantiated by the empirical evidence about belief formation in online contexts; see CR Sunstein, *Republic.com 2.0* (Princeton: Princeton University Press, 2007) [Sunstein].

This is especially the case if the speaker is a person in a position of authority or is perceived as having some kind of expertise. Speakers acquire authority or legitimacy through some kind of institutional support. This can take a number of different forms, including educational institutional support, religious institutional support, or state institutional support. On the basis of this institutional and social recognition of the speaker's authority, the speaker's speech acts have certain effects. The speech acts can enact norms, including norms that permit violence or norms that discriminate, or they can impose significant obligations on others to do what the speaker has advised or to accept what the speaker has said. The nature of the speaker's authority can make it difficult for listeners to rationally engage with what is said so as to form an independent judgment. This also includes those cases where it may be difficult to "speak back" to a person in an authoritative position because of the hearer's relative powerlessness. Finally, the socio-political climate in which the speech is made increases the likelihood of the speech act causing or constituting various harms. I have argued elsewhere that in these cases, authoritative speakers may be legally liable for harms caused or constituted by their speech acts, even if the speakers played no concrete role in bringing about the relevant harms.

Expert speakers differ from speakers with authority in the sense that they do not necessarily enact norms or impose obligations on others, but they are nevertheless able to secure uptake for their views in similar ways. People accept what these speakers say because of who they are and the perceived expertise that they have. While there is ongoing philosophical debate about the status of expert views and arguments from authority more generally, it is widely recognized and accepted that we often uncritically accept the views of experts and rely on them. As Ralph Johnson and J. Anthony Blair argue, the idea of an "autonomous, self-guiding, belief-scrutinizing individual as a person who sets off completely independently, an isolated rational agent," is a mistake.⁵⁴ In fact, we are quite dependent on others for much of what we know, because we do not have the time to investigate everything for ourselves, and much of what is accessible to us is made so through the testimony of others.⁵⁵

But how do ordinary persons judge whether someone is an expert when they themselves do not have the necessary expertise? Studies in expertise and experience suggest that in cases where direct assessment, internal to the expertise, is impossible, citizens can proceed indirectly, assessing the external features or signs of expertise the alleged experts display. According to Harry Collins and Robert Evans, this enables citizens "to make judgments between knowledge-claims based on something other than their scientific knowledge," specifically their "social knowledge."⁵⁶ Such assessments do "not depend on the understanding of the *expertise* being judged but upon an understanding of the *experts*."⁵⁷ In short, this means that ordinary citizens are often making "*social* judgments about *who* ought

⁵⁴ Cited in J Goodwin, "Accounting for the Appeal to the Authority of Experts" (2011) 25 *Argumentation* 287 [Goodwin].

⁵⁵ *Ibid.*

⁵⁶ H Collins & R Evans, *Rethinking Expertise* (Chicago: University of Chicago Press, 2007) at 56.

⁵⁷ *Ibid.*

to be agreed with, not *scientific* judgments about *what* ought to be believed.”⁵⁸ These social judgments can be facilitated or bolstered by checking the supporting documentation of the expert, like his or her institutional affiliation; CV; education; the opportunities he or she has had to develop the expertise; credentials; peer review of his or her work; track recognition; consistency with other experts; and consistency with other evidence.

Jean Goodwin sums up the basic logic of the appeal to authority as “believe this,’ an expert says, ‘because I say so.”⁵⁹ She elaborates:

[T]o actualize such an appeal, the expert must therefore make apparent to his audience of citizens *who* he is, and *that* he says so. The first task requires the expert to show that he is indeed an expert by offering his audience suitable tokens of expertise. This means that it is up to him to make available the sorts of signs that citizens can use to assess him: experiences, credentials, education, recognition and so on. The second task requires the expert to explicitly state his views as an expert on some matter of concern to his fellow citizens.⁶⁰

Goodwin argues that we should not be too concerned about speakers claiming expert knowledge because it enables us to locate responsibility and to hold certain speakers accountable for what they say.

However, speakers with expertise and authority need not always be legitimate to secure uptake for their views. Speakers can use the tropes and evidence of expertise or authority to give legitimacy or credibility to their views. These are cases of what I refer to as “manufactured authority.” Manufactured authority occurs in those cases where individuals may not have the kind of institutional support outlined above, or other legitimate supporting evidence of expertise, but are able to use the tropes or markers of authority and expertise to make it seem as though they do. Holocaust deniers attempt to manufacture authority for their views in precisely this way, and there is emerging evidence that some extreme right-wing groups are attempting something similar.

How do Holocaust deniers manufacture authority (and legitimacy) for their views? Some cases are straightforward in the sense that legitimate institutions, like tertiary institutions, affirm these views as subjects of serious academic debate. For example, Holocaust denial is considered a subject of serious intellectual discussion in the French academy, where deniers consider themselves to be pursuing an intellectually rigorous form of Holocaust revisionism. Some French university departments have even awarded postgraduate research degrees for Holocaust denial theses.⁶¹ Given the role of the intellectual in public life in France, this kind of institutional support is especially troubling.

However, even in cases where such institutional support is lacking, Holocaust deniers are able to seemingly manufacture legitimacy, and hence authority, by making their views seem as though they have serious academic merit. Most

⁵⁸ *Ibid.* at 47–48.

⁵⁹ J Goodwin, “Forms of Authority and the Real ad Verecundiam” (1988) 12 *Argumentation* 267–80, cited in Goodwin, *supra* note 54 at 292.

⁶⁰ Goodwin, *supra* note 54 at 292.

⁶¹ Fraser, *supra* note 36 at 522.

deniers claim to be historians, most are published authors, and most justify their speech on the grounds of free speech and other academic freedoms. As Catriona McKinnon has shown, many instances of Holocaust denial are presented as serious pieces of academic research and are thus the most dangerous forms of Holocaust denial. The Adelaide Institute founded by Toben, and the Institute for Historical Review (IHR) and its publication *The Journal of Historical Review* (JHR), are two such examples. The IHR statement of purpose claims that the Institute “is an independent educational research and publishing center that works to promote peace, understanding and justice through greater public awareness of the past, and especially socially-politically relevant aspects of twentieth-century history. We strive in particular to increase understanding of the causes, nature and consequences of war and conflict. We vigorously defend freedom of speech and freedom of historical inquiry.” The IHR’s journal is put together according to the conventions and style that govern genuine academic journals,⁶² and the website lists the conferences and speakers, many of them academics, hosted by the IHR. While the Adelaide Institute does not appear as polished at the IHR, it also has links to various expert authors who claim to be historians.

To the unknowing browser, the IHR and the Adelaide Institute can appear to be independent think tanks, and their publications can appear to be legitimate scholarly journals. Even though these organizations’ websites have no affiliation with any institution, it appears as though they do. As McKinnon notes, the majority of the papers that appear on the website deal with “revisionism” and the Holocaust, and none of the pieces have been critical of the Nazis’ anti-Jewish policies. With respect to the IHR, McKinnon writes: “[T]hose associated with the IHR—and those, such as Irving, who publish in the JHR—clearly conceive of themselves as talented historians whose political views are used to justify their exclusion from the Academy.”⁶³ Institutes like these attempt to manufacture legitimacy for the views they express by using the tropes or markers of expertise, and even though such legitimacy is not real, it appears as though it is. As such, the views can be deemed authoritative insofar as others might take the views to be the legitimate work of scholars. My point here is not that we should legally regulate this speech, or interpret hate speech regulations in a broad way so that these kinds of speech acts are captured. I do, however, think it is important to classify or categorize these speech acts in the right kind of way: they are not merely academic debate, but sophisticated forms of hate speech that might cause as much, if not more harm than obvious cases of extreme speech. It is relatively easy to see a racist rant for what it is, but it is not so easy to identify racism that looks like scholarship or serious (civil) political debate.

6. Objections

There are at least two different objections to the view I have been defending. The first is that these speakers and groups are on the fringes of political debate and so

⁶² C McKinnon, “Should We Tolerate Holocaust Denial?” 2007) 13:1 *Res Publica* 22.

⁶³ *Ibid.* at 23.

do not, in fact, have authority or expertise, manufactured or otherwise, in any degree that would seriously cause harm. For example, in his criticism of the *Keegstra* judgment, Terry Heinrichs discounts the authority Keegstra may have had, even though Heinrichs admits that the playing field in which speech is enacted is an inherently unequal one. Heinrichs states: “[A]re we seriously to believe that the ‘self-fulfillment’ of members of identifiable groups is threatened in any important way by the marginal mutterings of outsiders like Keegstra?”⁶⁴ Similarly, Moon points out that it is difficult to imagine that Keegstra’s bizarre views would be taken seriously by anyone who was not deeply mired in hatred, or who was not subordinate or in a vulnerable position in relation to the speaker.⁶⁵

The second objection is that my argument about uncritical audiences coming to accept these views because they think the speakers are credible assumes “the best about the speaker and the worst about his targets and his audience.”⁶⁶ That is, that the argument overestimates the talents of the hate speaker, and that it completely ignores the attributes of the audience and questions about the intelligence, moral character, education, upbringing, and intellectual acumen of the speaker’s audience. As Heinrichs puts it: “[A]nd so while it assumes that the former will be able to persuade his audience to accept whatever message he wishes to send, it also assumes that the latter will have little or no resistance to the message sent and will be unable to see the lack of moral character, socially destructive agenda, and intolerant motives of the speaker for what they are.”⁶⁷

In response to the first objection, while some persons in positions of authority might be nobodies to the wider community, they nevertheless hold significant sway over their audience. In any case, Keegstra was a high school teacher, and so it is reasonable to assume that his audience would have considered him to have some kind of expertise and authority. Authority does not mean authority over everyone; it simply means authority over a given domain. If one’s domain is the classroom, this is where the person will have the most influence; if a person is a religious leader, his authority is likely to be over his congregation. It does not matter that wider audiences do not accept the views and that they are capable of engaging with them; what matters is that persons within those domains are influenced by the views and are not in a position to critically engage with them. So it is not about underestimating general audiences, but about particular audiences coming to accept the norms.

In response to the second objection, it is not so much that audiences in general are uncritical, but that audiences who are already sympathetic to these views might not critically engage with them and may see them, instead, as expert validation, or that unknowing or uncritical audiences might come to accept the views because they consider them credible. Some audiences are critical, and some are not. This is borne out by the empirical studies carried out by Cass Sunstein and Lesley Wexler.

⁶⁴ T Heinrichs, “Censorship as Free Speech! Free Expression Values and the Logic of Silencing in *R v Keegstra*” (1997–8) 36:4 *Alta L Rev* 850.

⁶⁵ Moon, *supra* note 18 at 25.

⁶⁶ *Ibid.* at 870.

⁶⁷ *Ibid.*

These studies confirm that because the Internet facilitates contact with like-minded individuals, it often facilitates and strengthens fringe communities that have a common ideology, even though they are not geographically close. The Internet has made it possible for a diverse range of groups, including members of hate groups, to find each other, swap information, and encourage each others' beliefs. Sunstein notes: "[I]t is also clear that the Internet is playing a crucial role in permitting people who would otherwise feel isolated, or move on to something else, to band together and spread rumors, many of them paranoid and hateful."⁶⁸

Manufacturing authority and expertise in the ways outlined above can further strengthen existing convictions on the part of audiences, and it can also lead others to accept the views being disseminated. I am not suggesting that we ought to use legal mechanisms to regulate this kind of speech. The pragmatic difficulties with regulating online material, and the theoretical problems with censorship more generally are compelling enough arguments against regulation. What I am suggesting is that we should classify this speech as what it is: a form of sophisticated racism rather than academic debate or political discussion, and a form that might cause more harm than obvious instances of hate speech. If we are more attuned to the techniques and tropes used by extreme groups to manufacture legitimacy for their views, we might be better able to critically engage with such views, expose them for what they are, and so mitigate the harms they might cause.⁶⁹

Dr. Sarah Sorial
Senior Lecturer
Faculty of Law, Humanities and The Arts
The University of Wollongong
Wollongong, NSW 2522
sarahs@uow.edu.au

⁶⁸ Sunstein, *supra* note 53 at 58.

⁶⁹ I'd like to thank Yvonne Apolo for her research assistance. This project was supported by an Understanding Canada grant (2012), funded by the Canadian Government.