

Research Report - June 2020

Legal Aspects of Hate Speech in Canada

Lex Gill



3	About the Initiative
4	About the Author
5	Acknowledgements
6	Overview
7	Hate speech and the law in Canada
7	Key legislation provisions related to hate speech
10	Freedom of expression and “reasonable limits”
12	Restrictions on hate speech and the <i>Charter</i>
16	The definition of “hatred” in Canadian law
18	Emerging issues and new challenges
18	Hate speech and the internet
20	New recommendations before Parliament
23	Self-governance and heightened obligations for online platforms
26	Counterspeech and public education
28	Conclusion
29	Discussion Questions
33	Endnotes

About the Initiative



The Canadian Commission on Democratic Expression is a three-year initiative, led by the Public Policy Forum that aims to bring a concerted and disciplined review of the state of Canadian democracy and how it can be strengthened. The centerpiece is a small, deliberative Commission which will draw on available and original research, the insights of experts and the deliberations of a representative Citizen's Assembly to assess what to do about online harms and how to buttress the public good. The Commission is designed to offer insights and policy options on an annual basis that support the cause of Canada's democracy and social cohesion. The Commission is supported by national citizen assemblies as well as by an independent research program.

This initiative grew out of earlier insights about the relationship of digital technologies to Canada's democracy covered by the Public Policy Forum's ground-breaking report, *The Shattered Mirror* and its subsequent interdisciplinary research outlined in the *Democracy Divided* report (with UBC) and through the Digital Democracy Project partnership with McGill university.

The initiative is stewarded by Executive Director, Michel Cormier and delivered in partnership with MASS LBP and the Centre for Media, Technology and Democracy at McGill University's Max Bell School of Public Policy, who are executing the national citizen assemblies and research program, respectively.

To learn more about the initiative and how you can become involved, please visit www.ppforum.ca. The initiative will run from April 2020 to March 2023.

This project has been made possible in part by the Government of Canada. PPF would also like to thank the McConnell Foundation for their support.

About the Author

Lex Gill

Lawyer, Trudel Johnson & Lespérance (TJL)
& Legal Fellow at Citizen Lab

M^e Lex Gill is a lawyer with degrees in civil and common law from McGill University. She is currently employed at Montréal's Trudel Johnston & Lespérance (TJL), a firm known for class actions and groundbreaking public interest litigation in areas like human rights, environmental law, and corporate accountability. Lex has been recognized for her work at the intersection of constitutional law, emerging technology, and civil liberties, and has published influential work in these areas. Prior to joining TJL, she clerked for Chief Justice Richard Wagner of the Supreme Court of Canada, worked at Citizen Lab, the Canadian Civil Liberties Association, the Canadian Internet Policy and Public Interest Clinic, and Harvard's Berkman Klein Center. She has served on numerous advisory boards and has won several awards for her academic work and commitment to social justice. Alongside her practice, she is teaching a course on technology and the *Charter* at McGill's Faculty of Law.

The author would like to thank Meghan Sali (University of Ottawa, Faculty of Law) for substantive legal research and editing in support of this report, Cynthia Khoo and Tamir Israel for their thoughtful and detailed feedback on an earlier draft, and Sonja Solomun and Taylor Owen for the opportunity to contribute to this important initiative.

Overview

This brief report was prepared for the Canadian Commission on Democratic Expression (CCDE) in March 2020. It aims to support the Commission's efforts to better understand hate speech in Canada and to respond through research, public consultation, stakeholder outreach and other forms of democratic engagement.

It begins by providing an overview of hate speech and its relationship to freedom of expression in Canadian law. It explores the legal tensions and policy questions that have historically animated debates surrounding hate speech in Canada, and summarizes some of the most significant recent developments. The report also examines some of the ways that technological, economic and social change have begun to challenge the ways we think about harmful expression.

Hate speech is a topic that is sometimes polarizing or divisive. This report does not attempt to canvass every potential view on the subject, and is only intended as a starting point for further dialogue. At the end of the document, readers will find a list of discussion questions meant to encourage reflection, reveal common ground and provoke new insights within their own communities.

The words we choose to describe social harms have legal implications as well as symbolic and political power.¹ This report generally uses the language of “hate speech” (which tends to be the common term used by Canadian courts), or “hate propaganda” when describing the *Criminal Code* offence. However, readers will note that certain authors choose different language, some of which may encompass expression that is lawful in Canada (e.g., “harmful speech”²), while others suggest a more narrow scope than the Canadian legal definition of hate speech would tend to include (e.g., “violent” or “dangerous” speech³).

Hate speech and the law in Canada

Key legislative provisions related to hate speech

In Canada, hate speech is subject to both criminal and administrative sanctions. In the criminal context, section 319(2) of the *Criminal Code* states that “every one who, by communicating statements, other than in private conversation, wilfully promotes hatred against any identifiable group” is guilty of an offence.⁴ In this context, an identifiable group “means any section of the public distinguished by colour, race, religion, national or ethnic origin, age, sex, sexual orientation, gender identity or expression, or mental or physical disability.”^{5,6} Because section 319(2) is a criminal law, the prosecution is required to prove each element of the offence beyond a reasonable doubt. Prosecutions under the *Criminal Code* for promotion of hatred require the approval of the Attorney General to proceed, a rule which may partly explain why such prosecutions have been rare.⁷

There is notably no legal requirement to prove that the communication actually *caused* hatred or harm.^{8,9} However, a conviction under s. 319(2) does require a finding that the accused person “wilfully promoted” hatred against a group. The Ontario Court of Appeal has confirmed that this subjective intent requirement can include willful blindness to the fact that the promotion of hatred would be a substantially certain consequence of one’s communications.^{10,11,12} The provision also lists four specific defences,¹³ noting that an individual cannot be convicted under the following conditions:

- a. if he establishes that the statements communicated were true;¹⁴

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

Section 319(2) is closely related to several other provisions of the *Criminal Code*, including section 318, which criminalizes advocating or promoting genocide;^{15,16} section 319(1), which prohibits publicly communicating statements that incite hatred against an identifiable group where that incitement is likely to lead to a breach of the peace; and section 430(4.1), which prohibits mischief in relation to property that is motivated by bias, prejudice or hate based on colour, race, religion, national or ethnic origin, age, sex, sexual orientation, gender identity or expression or mental or physical disability.¹⁷

Furthermore, where a criminal offence (such as assault or harassment) is “motivated by bias, prejudice or hate based on race, national or ethnic origin, language, colour, religion, sex, age, mental or physical disability, sexual orientation, or gender identity or expression, or on any other similar factor” that fact is considered an aggravating factor at the sentencing stage.^{18,19} In 2018, there were 1,798 such crimes reported by police in Canada, the vast majority of which were motivated on the basis of race, ethnicity, religion or sexual orientation specifically.^{20,21} Statistics Canada indicates that 364 police-reported hate crimes were also categorized by law enforcement as “cybercrimes” between 2010 and 2017, and that the groups most commonly targeted by these crimes were Muslim (17%), LGBTQ2 (15%), Jewish (14%) and black (10%).²² It is important to understand that these numbers are lower than the actual rate of hate-related offences: an estimated two out of three victims do not report to police at all.²³ Systemic mistreatment and discrimination by law enforcement means that individuals who belong to some marginalized or disadvantaged groups are notably less likely to report to police, despite the fact that they are the primary targets of crimes motivated by hatred.²⁴

Several provincial and territorial human rights statutes also contain provisions prohibiting the promotion of hatred.²⁵ However, and unlike the *Criminal Code* provisions discussed above, these laws give claimants the right to administrative law remedies. As noted by the Supreme Court in *Canada (Human Rights Commission) v Taylor*, the aim of human rights legislation “is not to bring the full force of the state’s power against a blameworthy individual for the purpose of imposing punishment. Instead, [such provisions] generally operate in a less confrontational manner, allowing for a conciliatory settlement if possible and, where discrimination exists, gearing remedial responses more towards compensating the victim.”^{26,27} Because these provisions are not accompanied by criminal penalties and serve a different social purpose, they may also capture a broader range of expression and behaviour. For example, given that human rights legislation tends to focus on the discriminatory *effects* of harmful conduct, it does not generally require subjective intent on the part of the communicating party.^{28,29} Human rights laws have a “quasi-constitutional” character in Canada, are afforded a large and liberal interpretation and have primacy over other statutes.³⁰

Until its repeal in 2013, section 13 of the *Canadian Human Rights Act* also gave rise to administrative law remedies for hate speech at the federal level. In particular, section 13 of the *Act* declared it a discriminatory practice to “communicate telephonically,” including through the use of the internet or similar means, “any matter that is likely to expose a person or persons to hatred or contempt by reason of the fact that that person or those persons are identifiable on the basis of a prohibited ground of discrimination.”³¹ Though section 13 had been the subject of significant criticism (notably, but not exclusively, from civil liberties groups), its repeal was also controversial.³² As discussed below, several constituencies continue to support its reintroduction in either full or amended form.³³

Protections against hate speech also have a basis in international law. Most notably, Article 20(2) of the *International Covenant on Civil and Political Rights*, to which Canada is a party, states that “any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.”³⁴ Similarly, Article 4(a) of the *International Covenant on the Elimination of All Forms of Racial Discrimination* requires State Parties to “declare an offence punishable by law all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin, and also the provision of any assistance to racist activities, including the financing thereof.”³⁵

Freedom of expression and “reasonable limits”

When the government adopts legislation or takes action to restrict hate speech, it imposes a limit on freedom of expression, a right which is constitutionally protected under the *Canadian Charter of Rights and Freedoms*.³⁶ While the Charter provides the legal foundation for almost any discussion related to hate speech in Canada, it is important to note that freedom of expression, belief and opinion are also protected under the Canadian Bill of Rights³⁷ (a federal human rights statute that preceded the Charter), Quebec’s Charter of Human Rights and Freedoms,³⁸ and various other provincial and federal statutes.³⁹ These rights are also enshrined in several international law instruments, including Article 19 of the International Covenant on Civil and Political Rights⁴⁰ and Article 19 of the Universal Declaration of Human Rights.⁴¹

The Charter safeguards “freedom of thought, belief, opinion and expression” under section 2(b), which also includes “freedom of the press and other media of communication.”^{42,43} This guarantee is listed among several other “fundamental freedoms,” which include freedom of conscience, religion, peaceful assembly and association.⁴⁴ The Constitution recognizes that these rights belong to all individuals.

Courts have found that section 2(b) affords constitutional protection to all forms of expressive content—that is, “any activity or communication that conveys or attempts to convey meaning.”^{45,46} The scope of “expression” under the provision is therefore very broad, and encompasses everything from picketing and political protest to child pornography and defamatory libel.⁴⁷ Section 2(b) is also sometimes described as “content-neutral,” in the sense that it applies to all expressive content, even where that expression may be false, unpopular, disturbing or offensive.^{48,49,50} In this sense, hate speech is therefore considered protected expression under section 2(b) of the *Charter*.

However, constitutional protection for freedom of expression is neither absolute nor unlimited in Canada. Section 1 of the *Charter* provides that the rights and freedoms guaranteed therein can be subject to certain “reasonable limits” by the state—but only where those limits “can be demonstrably justified in a free and democratic society.”⁵¹ Ultimately, this requires courts to conduct a proportionality analysis that considers the importance of the government’s objectives in

restricting the expression, whether that restriction is rationally connected to those objectives, whether the limit is sufficiently tailored such that it impairs the right as little as reasonably possible, and whether the negative impacts of the measure outweigh its beneficial effects.⁵²

In other words, whether or not a measure that restricts freedom of expression is found to be constitutional will usually come down to a balancing exercise. This does not mean that the analysis is entirely subjective—courts hear submissions on these issues from a wide variety of actors and are bound by extensive jurisprudence on the issue. However, it does mean that there is no perfect calculus for determining the constitutional limits of speech in Canada. The analysis is inherently contextual, and necessarily raises complex questions about the limits of state power, what it means to live in a free society, and whose voices and interests deserve protection.

In determining whether a given restriction on freedom of expression represents a “reasonable limit” in the constitutional sense, Canadian courts have tended to root their analysis in the underlying purposes of section 2(b) of the *Charter*. In other words, the analysis begins from the perspective that there are certain reasons that freedom of expression is protected in a democratic society.

The Supreme Court has generally offered three such reasons.⁵³ The first is that freedom of expression allows individuals to fully participate in social and political decision-making, and is therefore essential to democratic self-governance. Second, there is the view that freedom of expression promotes the search for and the attainment of truth, through what has sometimes been described as “a competitive marketplace of ideas.”⁵⁴ Third, there is the view that freedom of expression is worthy of protection due to its inherent capacity to facilitate self-discovery, self-fulfillment and human flourishing. It is important to remember that these values—while fundamental to the liberal democratic tradition—reflect certain assumptions about human behaviour and political life.

Nonetheless, Canadian constitutional jurisprudence has been clear that when the expression at issue is closely connected to one or more of these core values, state-imposed limits on that expression will be more difficult to justify. Conversely, where the expression in question is only peripherally or marginally connected to these values, it will be easier for the state to justify a restriction. This is part of the reason why tobacco companies’ rights to advertise cigarettes to young people is afforded less constitutional protection than a student group’s right to encourage youth to vote.⁵⁵

Restrictions on hate speech and the *Charter*

Courts have used these constitutional principles both to define hate speech in Canada and to evaluate the constitutionality of government restrictions on that speech. Some of the most significant developments in this area arise from a trilogy of Supreme Court cases released in the early 1990s. These three decisions—*R v Keegstra*, *R v Andrews*, and *Taylor*—affirmed the constitutionality of prohibitions on hate speech in the context of both criminal and human rights law,^{56,57,58} including section 319(2) of the *Criminal Code* and former section 13(1) of the *Canadian Human Rights Act* (CHRA), both discussed above. Drawing on international law, philosophy and social science evidence, these cases continue to provide the foundation for the Canadian legal approach to hate speech.

In the trilogy, the Court made it clear that hate speech was protected by the freedom of expression guarantee of the *Charter*. In *Keegstra*, as in *Taylor*,⁵⁹ the Court specifically refused to deem hate speech analogous to “violence”—that is, expression communicated directly through physical harm^{60,61}—which would have excluded it from section 2(b)’s protection altogether. However, majorities of the Court found that the limits imposed on hate speech through the laws in question were justified as “reasonable limits” under section 1.

It is notable that the *Keegstra* majority distanced itself from U.S. First Amendment jurisprudence, which had generally found the suppression of hate speech to be incompatible with freedom of expression. The majority observed that “the international commitment to eradicate hate propaganda and, most importantly, the special role given equality and multiculturalism in the Canadian Constitution necessitate a departure from th[is] view.”⁶² Instead, the Supreme Court’s analysis during this period was largely anchored in the work of the 1966 *Special Committee on Hate Propaganda in Canada* (the Cohen Committee), which had led to Parliament’s adoption of the *Criminal Code* provisions at issue in *Keegstra* and *Andrews*. The preamble of the Committee’s report read as follows:

This Report is a study in the power of words to maim, and what it is that a civilized society can do about it. Not every abuse of human communication can or should be controlled by law or custom. But every society from time to time draws lines at the point where the intolerable and

the impermissible coincide. In a free society such as our own, where the privilege of speech can induce ideas that may change the very order itself, there is a bias weighted heavily in favour of the maximum of rhetoric whatever the cost and consequences. But that bias stops this side of injury to the community itself and to individual members or identifiable groups innocently caught in verbal cross-fire that goes beyond legitimate debate.⁶³

Informed by the Cohen Committee's work, the analysis in these early cases centred on the serious harms that Parliament had sought to prevent by enacting section 319(2) of the *Criminal Code* and section 13 of the *CHRA*. In the view of the *Keegstra* majority, government measures to restrict hate propaganda addressed two principle types of harm. First, they sought to prevent the personal humiliation and degradation caused by hate propaganda, stating that "the emotional damage caused by words may be of grave psychological and social consequence."⁶⁴ In this respect, the Court understood criminal sanctions for hate propaganda as protecting individual human dignity. In a similar vein, the Taylor majority acknowledged the foreseeable harm caused by "loss of self-esteem, feelings of anger and outrage and strong pressure to renounce cultural differences that mark [individuals subjected to racial or religious hatred] as distinct."^{65,66}

The second type of harm concerns the effects of hate propaganda on society more broadly, including the risk that failure to act could desensitize the public to escalating forms of persecution or give proponents of that speech license to commit acts of discrimination and violence. Over two decades later in *Saskatchewan (Human Rights Commission) v Whatcott*, the Court summarized the issue lucidly:

Hate speech, therefore, rises beyond causing emotional distress to individual group members. It can have a societal impact. If a group of people are considered inferior, subhuman, or lawless, it is easier to justify denying the group and its members equal rights or status. As observed by this Court in *Mugesera*, the findings in *Keegstra* suggest "that hate speech always denies fundamental rights." As the majority becomes desensitized by the effects of hate speech, the concern is that some members of society will demonstrate their rejection of the vulnerable group through conduct. Hate speech lays the

groundwork for later, broad attacks on vulnerable groups. These attacks can range from discrimination, to ostracism, segregation, deportation, violence and, in the most extreme cases, to genocide.⁶⁷

When *Keegstra* was heard, the perception that hate speech could be effectively countered through public debate and greater access to information had been losing ground. Instead, the majority echoed the trepidation expressed by the Cohen Committee when it wrote that:

...we are less confident in the 20th century that the critical faculties of individuals will be brought to bear on the speech and writing which is directed at them. In the 18th and 19th centuries, there was a widespread belief that man was a rational creature, and that if his mind was trained and liberated from superstition by education, he would always distinguish truth from falsehood, good from evil. ... We cannot share this faith today in such a simple form. While holding that over the long run, the human mind is repelled by blatant falsehood and seeks the good, it is too often true, in the short run, that emotion displaces reason and individuals perversely reject the demonstrations of truth put before them and forsake the good they know. The successes of modern advertising, the triumphs of impudent propaganda such as Hitler's, have qualified sharply our belief in the rationality of man.⁶⁸

Though these words were published in 1966 and endorsed by the Supreme Court in 1990, they remain remarkably evocative of modern debates surrounding hate speech, misinformation and propaganda online. They call into question liberal assumptions about the “marketplace of ideas” that underpinned earlier free speech discourse, and reject the view that all speech is equally valuable in a democratic society.⁶⁹

Indeed, the majority in *Keegstra* was explicit in its finding that the *Criminal Code* provisions affected a narrow and circumscribed category of expression that was “only tenuously connected with the values underlying the guarantee of freedom of speech.”⁷⁰ Though it acknowledged that hate propaganda was a kind of political

speech, and “thus putatively ... at the very heart of the principle extolling freedom of expression as vital to the democratic process,” it concluded that such speech actually undermined and subverted democratic ideals by denying respect and dignity to individuals on the basis of their identity. This was echoed by the Court in the 2013 *Whatcott* decision, in which hate speech was described as an “extreme and marginal type of expression” that could contribute little to the values underlying section 2(b).⁷¹

The dissent in *Keegstra*, by contrast, would have found that section 319(2) of the *Criminal Code* was not constitutional and that it did not constitute a reasonable limit on freedom of expression under section 1 of the *Charter*. In its view, the government’s objective and the specific criminal law measures prohibiting hate propaganda were only tenuously connected. It also found that the legislative measures were neither minimally impairing of the right to free expression nor proportionate in the circumstances. Among the many concerns expressed by the dissenting judges, the risk that the language of section 319(2) was so broad and ill-defined that it risked “a chilling effect on defensible expression by law-abiding citizens” was paramount.⁷² In a passage regarding the overbreadth of the provision, the dissent noted that section 319(2) had already “provoked many questionable actions on the part of the authorities” to target unpopular groups, including decisions to censor controversial books and films at the border, and the arrest of anti-American pamphleteers.⁷³ Indeed, the possibility that hate speech laws might be abused continues to concern prominent human rights advocates. As the UN Special Rapporteur on freedom of expression David Kaye recently wrote in the international law context:

[The term’s] vagueness and the lack of consensus around its meaning can be abused to enable infringements on a wide range of lawful expression. Many Governments use “hate speech”, similar to the way in which they use “fake news”, to attack political enemies, non-believers, dissenters and critics.⁷⁴

The concerns expressed by the dissent in *Keegstra* and *Taylor* regarding the potential chilling effects of hate speech legislation are particularly interesting in light of the Court’s reasoning in *Whatcott* two decades later, where it found that failure to prevent the proliferation of hate speech had the effect of silencing the voices of individuals and groups targeted by hate speech:

Indeed, a particularly insidious aspect of hate speech is that it acts to cut off any path of reply by the group under attack. It does this not only by attempting to marginalize the group so that their reply will be ignored: it also forces the group to argue for their basic humanity or social standing, as a precondition to participating in the deliberative aspects of our democracy.⁷⁵

The dissent in *Keegstra* also expressed concern that prosecutions under the *Code* threatened to turn accused persons into martyrs or inadvertently provoke sympathy and belief: “theories of a grand conspiracy between government and elements of society wrongly perceived as malevolent,” they noted, “can become all too appealing if government dignifies them by completely suppressing their utterance.”^{76,77} In their view, human rights legislation—which offers the possibility for reparation without the stigma or harsh sanctions (including the threat of imprisonment) of a criminal sentence—tended to provide a more appropriate avenue for redress.⁷⁸ Nonetheless, the dissenting judges in *Keegstra* would have also found section 13 of the *CHRA*’s administrative scheme unconstitutional for many of the same reasons expressed in *Taylor*, arguing that the provision was ill-defined, overbroad and lacked the defences necessary to safeguard lawful, democratic and truthful speech.⁷⁹

The definition of “hatred” in Canadian law

It is important to understand that the concept of “hatred” (whether in the criminal law or human rights context) has been interpreted fairly strictly in Canada in order to avoid encroaching upon lawful speech that is merely distasteful, upsetting or offensive. For example, in *Taylor* the majority emphasized that the “hatred or contempt” described in former section 13 of the *CHRA* referred only to “unusually strong and deep-felt emotions of detestation, calumny and vilification” that were “ardent and extreme” in nature.⁸⁰ Similarly, in *Keegstra*, the majority found that the term connotes “a most extreme emotion that belies reason ... [that] if exercised against members of an identifiable group, implies that those individuals are to be despised, scorned, denied respect and made subject to ill-treatment on the basis of group affiliation.” In the Ontario Court of Appeal’s decision in *Andrews*, Cory J.A.

wrote that hatred is that which “instil[s] detestation” and “lays the foundations for the mistreatment of members of the victimized group.”⁸¹

Canadian law considers the concept of hatred to go far beyond hurtful or distasteful words; instead, it is an emotion that calls into question the basic humanity of the other.⁸² In *Keegstra*, the majority cautioned judges to offer specific jury instruction regarding this aspect of section 319(2), requiring guidance that “include[s] express mention of the need to avoid finding that the accused intended to promote hatred merely because the expression is distasteful.”⁸³ The concept of “hatred” was also explicitly distinguished from other sentiments in *Whatcott*, a case in which the Court struck the words “ridicules, belittles or otherwise affronts the dignity of” from section 14(1)(b) of the Saskatchewan Human Rights Code on the basis that they constituted an impermissible encroachment on freedom of expression:

In my view, expression that “ridicules, belittles or otherwise affronts the dignity of” does not rise to the level of ardent and extreme feelings that were found essential to the constitutionality of s. 13(1) of the *CHRA* in *Taylor*. Those words are not synonymous with “hatred” or “contempt”. Rather, they refer to expression which is derogatory and insensitive, such as representations criticizing or making fun of protected groups on the basis of their commonly shared characteristics and practices, or on stereotypes.⁸⁴

The Court maintained, however, that the law could continue to restrict speech which exposes a person or class of persons to actual hatred, by asking whether, “in the view of a reasonable person aware of the context and circumstances,” the representation exposes or tends to expose those persons “to detestation and vilification on the basis of a prohibited ground of discrimination.”⁸⁵

Emerging issues and new challenges

The 1990 trilogy, and the body of jurisprudence that followed, have provided a relatively solid foundation for lower courts and human rights tribunals to apply hate speech laws in Canada. However, the approach taken by Canadian jurists is not without critics. Some authors have asserted that Canadian courts have been insufficiently protective of 2(b) rights, while others have argued that they have not gone far enough to defend the rights and dignity of marginalized groups. Even so, the constitutionality of the *Criminal Code* offences concerning hate propaganda has not been seriously called into question since *Keegstra* and *Andrews*, despite an attempt to do so by a man convicted of willfully promoting hatred against Jewish people in British Columbia in 2015.⁸⁶ Though political debates surrounding the appropriate limits of free expression remain a persistent feature of the country's democratic life, hate speech remains a relatively stable concept in Canadian law. With this in mind, the following sections of the report explore a few recent developments and emerging challenges that may be of potential interest to the Commission.

Hate speech and the internet

It is difficult to find accurate and complete statistics regarding incidents of online hate speech; variations in methodology, time period, definition and platform mean that data tends to be either over-inclusive or under-inclusive. Statistics can nonetheless offer some insight regarding the scope of the phenomenon. For example, between July and September 2019, Facebook reported that it took action

on 7 million pieces of hate speech, up from 2.9 million for the same quarter in 2018.⁸⁷ During that same period, YouTube removed over 500 million comments from its platform, 5.8% of which were deleted on the basis that they had been deemed “hateful or abusive.”⁸⁸ Notably, in the following quarter (from October to December 2019), 24.7% of all content removed by YouTube was deleted on the basis that it fell within this category⁸⁹—a significant shift that appears to reflect the platform’s changing priorities.⁹⁰

Users’ self-reported experiences provide another useful data point. For example, a recent study indicates that a third of internet users in Canada encounter content that they would consider hate speech at least once a week online, though it is not clear how much of that content corresponds to the legal definition of hate speech in Canada. It is worth noting that among the subset of survey respondents who identified as belonging to a minority group in that study, the number who reported encountering hate speech rose to 44%.⁹¹ An expansive report conducted by the Commission des droits de la personne et des droits de la jeunesse concerning hate crimes in Quebec found that a third of participants had been personally subjected to xenophobic or Islamophobic insults or threats online.^{92,93} These findings correspond with the reality that marginalized groups are disproportionately targeted by abusive comments online: for example, a study of 70 million comments left on the *Guardian* news website between 2006 and 2016 revealed that of the 10 most abused writers, eight were women and the two men were black.^{94,95}

Over the last two decades, Canadian courts, administrative bodies and legal scholars have given considerable thought to the issue of hate speech online. Just months before the law repealing section 13 of the *CHRA* came into effect in 2014,⁹⁶ the Federal Court of Appeal issued reasons in a case called *Lemire*, in which it had been argued that restricting communication of hate messages over the internet under the Act constituted an unconstitutional breach of *Charter*-protected rights.⁹⁷ In the Court of Appeal’s unanimous view, the mere fact that impugned speech took place over the internet was insufficient to render the provision unconstitutional or to depart from the Supreme Court’s conclusions in *Taylor*.⁹⁸ “Indeed,” the Court noted, “a statutory prohibition of the communication of hate speech without including such a widely used and powerful means of communication as the internet would be an exercise bordering on futility.”⁹⁹ Notably, it had been argued that the existence of a complaint under section 13 could pressure internet service providers to block lawful content or ban users preemptively and before an actual violation was found, but the Court found that this risk was insufficient to render the law itself unconstitutional.¹⁰⁰

More recently, the federal government announced 10 principles meant to guide Canadian digital policy under the banner of a “Digital Charter,” which included one principle asserting that “Canadians can expect that digital platforms will not foster or disseminate hate, violent extremism or criminal content.”¹⁰¹ The Digital Charter background materials link this principle to Canada’s recent decision to sign the Christchurch Call to Action¹⁰²—a set of voluntary commitments on the part of governments and internet service providers to counter the proliferation of “terrorist and extremist content” online.¹⁰³

New recommendations before Parliament

In June 2019, the House of Commons Standing Committee on Justice and Human Rights (JUST) released a 72-page report called “Taking Action to End Online Hate,” which was the result of a consultation process that included a number of Canadian civil society organizations.

The study was conducted in response to groups that had raised alarm regarding a 2017 spike in police-reported hate crimes, and presumes a clear causal link between such crimes and online hate speech.^{104, 105, 106} Before the Committee, a number of groups urged measures that would allow more aggressive use of the *Criminal Code* provisions targeting the promotion of hatred, including through increased resources for law enforcement, greater international collaboration in the prosecution of hate crimes, and repeal of the provision requiring the Attorney General’s consent to institute proceedings.¹⁰⁷ The Committee also received submissions that the concept of “hatred” needed to be better defined¹⁰⁸ or better publicly understood.¹⁰⁹ It also heard several calls for a national strategy to counter online hate.¹¹⁰

The Committee also heard evidence from several witnesses expressing concern regarding the lack of reliable data on hate crimes and hate incidents, including online hate speech. The proposed responses ranged from better tools for individual reporting on social media platforms to increased data collection by technology companies, in partnership with the state.¹¹¹

The report concludes that self-regulation and user-initiated reporting common to most major online platforms did not provide an adequate response to online hate speech. Instead, the Committee found that “government leadership is necessary to regulate social media companies” and that a rights-based regulatory framework should be developed in collaboration with civil society.¹¹² While some witnesses expressed a desire for platforms to take more immediate action to remove harmful content, others raised the concern that harsh penalties could encourage platforms to favour overbroad enforcement measures and censorship for the sake of convenience.¹¹³ Some witnesses suggested that a government body be entrusted with a new oversight and enforcement role with regard to online platforms,¹¹⁴ while others reiterated that even well-intended measures to curtail hate speech online risked jeopardizing civil liberties.¹¹⁵

The Committee ultimately adopted nine recommendations in response to these submissions, including to increase justice system funding, invest in public education and digital literacy, develop best practices for law enforcement and adopt strategies for better data collection and reporting on hate crimes and hate incidents, including the development of a national database. These proposals are somewhat aligned with the Broadcasting and Telecommunications Legislative Review Panel’s recent recommendation that the federal government periodically “review the efficiency of enforcement mechanisms for monitoring and removing illegal content and conduct found online” in collaboration with provincial and territorial governments.¹¹⁶ The Committee also called upon the Government of Canada to formulate a definition of “hate” or “hatred” in line with Supreme Court jurisprudence that “acknowledges persons who are disproportionately targeted by hate speech including but not limited to racial, Indigenous, ethnic, linguistic, sexual orientation, gender identity, and religious groups.”

Several of the Committee’s recommendations would require legislative reform to enact. The first is to provide a civil remedy to fill the perceived gap created by the repeal of former section 13 of the *CHRA*. The report states that the remedy “could take the form of reinstating the former section 13 of the *Canadian Human Rights Act*, or implementing a provision analogous to the previous section 13 within the *Canadian Human Rights Act*, which accounts for the prevalence of hatred on social media.”¹¹⁷

Second, the Committee made the following recommendation:

That the Government of Canada establish requirements for online platforms and Internet service providers with

regards to how they monitor and address incidents of hate speech, and the need to remove all posts that would constitute online hatred in a timely manner.

These requirements should set common standards with regards to making reporting mechanisms on social media platforms more readily accessible and visible to users, by ensuring that these mechanisms are simple and transparent.

Online platforms must have a duty to report regularly to users on data regarding online hate incidents (how many incidents were reported, what actions were taken/what content was removed, and how quickly the action was taken). Failure to properly report on online hate, must lead to significant monetary penalties for the online platform.

Furthermore, online platforms must make it simple for users to flag problematic content and provide timely feedback to them relevant to such action.

Almost all large social media platforms have voluntarily adopted policies addressing hate speech and harmful content.¹¹⁸ However, the proposal advanced by the Committee would require legislative reform to impose mandatory obligations for content monitoring and removal, depending on what is meant by requirements “with regards to... the need to remove all posts that would constitute online hatred in a timely manner”.

As with almost all efforts to regulate foreign technology companies, the proposed measures could provoke complex questions regarding jurisdiction and enforcement. In the event that Parliament enacted such a law, some of the powers described may also be vulnerable to constitutional challenge as a violation of section 2(b) or section 8 privacy rights under the *Charter*. The UN Special Rapporteur on freedom of expression has recently provided extensive guidance to States considering such measures, noting their potential to interfere with international human rights law and the need for extensive procedural safeguards.¹¹⁹ It is worth noting that Supreme Court jurisprudence sets a high bar for injunctive relief in the context of hate speech.¹²⁰

Self-governance and heightened obligations for online platforms

Technology companies and online platforms play an outsized role in shaping what freedom of expression, democratic participation, privacy and substantive equality look like in the 21st century. It is without question that these companies have a special responsibility to protect and respect human rights.¹²¹ While these commercial actors have increased voluntary efforts to address issues like hate speech in recent years,¹²² they have not always lived up to their obligations in practice. Facebook’s recent admission that it failed to intervene when civil society actors warned that its platform was being systematically abused to fuel a campaign of ethnic cleansing against Myanmar’s Rohingya Muslim minority offers just one stark and devastating example.^{123,124}

At the same time, attempts to regulate these companies by imposing legal obligations to monitor and remove certain kinds of content tend to raise serious constitutional and international human rights issues. Citizen Lab, a leading research organization in this area, has consistently demonstrated that technology designed to filter harmful content online—even when the original purpose of the measure is beneficial, lawful or relatively innocuous—can invite serious abuse by private actors and governments alike.^{125,126} For example, filtering technology developed by a Canadian company called Netsweeper has been used to facilitate censorship of political content, LGBTQ2 websites (including healthcare and HIV-related information) and independent journalism in countries like the United Arab Emirates, Bahrain, Yemen and beyond.¹²⁷ Research also demonstrates that information controls tend to tighten around politically sensitive events, such as elections in Iran¹²⁸ or the COVID-19 outbreak in China,¹²⁹ necessitating even greater caution when considering the adoption of technology that facilitates the surveillance and censorship of online expression.

Requiring platforms or internet service providers to automatically and preemptively remove harmful content also implies their ability to rapidly identify that content. In this sense, it is important for readers to understand that content filtering and censorship

technology is almost always surveillance technology as well. It is therefore rare

Requiring platforms or internet service providers to automatically and preemptively remove harmful content also implies their ability to rapidly identify that content.

that the adoption of such measures will not involve at least indirect impacts on users' privacy rights.^{130,131} Groups and individuals subject to persecution on the basis of identity or political affiliation are often the first to be targeted for censorship and surveillance online, and the real-life consequences for these individuals and their families can be grave. According to Amnesty International, in 2016 people were arrested for what they said online in 55 countries worldwide.¹³²

While there is some enthusiasm for greater use of automation, machine learning and artificial intelligence to help identify and remove hate speech online, determining what constitutes hate speech and how it weighs against countervailing freedom of expression concerns is an intensely contextual (and therefore resource-intensive) exercise.¹³³ Even with the most advanced technologies, conversations between people promoting hatred and people targeted by it can look extremely similar, creating a risk that the speech of marginalized groups will be inadvertently censored.¹³⁴

In response to the risk of overbroad or counterproductive measures, some researchers and companies have experimented with strategies that may be seen as a lesser infringement of the rights at stake—for example, by using automated systems to “quarantine” comments flagged as potential hate speech prior to human review¹³⁵ or by limiting the speed at which certain kinds of messages can be forwarded.¹³⁶ UN Special Rapporteur on freedom of expression David Kaye has noted that most technology companies have access to a wide variety of tools in response to hate speech beyond outright removal, a fact which may eventually have implications for the section 1 analysis in Canada:

They can delete content, restrict its virality, label its origin, suspend the relevant user, suspend the organization sponsoring the content, develop ratings to highlight a person's use of prohibited content, temporarily restrict content while a team is conducting a review, preclude users from monetizing their content, create friction in the sharing of content, affix warnings and labels to content, provide individuals with greater capacity to block other users, minimize the amplification of the content, interfere with bots and coordinated online mob behaviour, adopt geolocated restrictions and even promote counter-messaging. Not all of these tools are appropriate in every circumstance, and they may require limitations themselves,

but they show the range of options short of deletion that may be available to companies in given situations.¹³⁷

It must be conceded that the financial incentives for these companies to filter and remove hate speech in a rights-protective manner that responds to context are far from obvious—particularly in the absence of legal or financial consequences for the wrongful removal of lawful expression. Moreover, by imposing these kinds of obligations in the absence of a framework for independent oversight and review (whether by courts or similar institutions) governments may inadvertently delegate the power to determine the limits of lawful expression to private commercial actors. Because the appropriate legal limits on freedom of expression vary considerably from one jurisdiction to the next—even among democracies with highly developed legal systems—enforcement of Canadian legal rules online can also raise tricky questions about jurisdiction, extraterritorial enforcement and comity.¹³⁸

Indeed, hate speech, whether online or off, is not a uniquely Canadian problem, and is the subject of advocacy, scholarship and complex policy debates across a wide variety of cultures and legal contexts.¹³⁹ Different legal and constitutional traditions have developed a variety of approaches to the treatment of hate speech, and some states are increasingly seeking to enlist online platforms in the enforcement of those laws.

For example, the recently adopted Network Enforcement Act in Germany creates legal requirements for online platforms to remove content considered illegal under the German Criminal Code, including provisions criminalizing hate speech, and imposes significant fines on technology companies that fail to comply.¹⁴⁰ Similar legislation has been proposed in France.¹⁴¹ While the Network Enforcement Act appears, at least preliminarily, to have resulted in a significantly higher volume of content takedowns by technology companies,¹⁴² it is much less clear what proportion of the deleted content is unlawful in fact. Indeed, the history of copyright enforcement under the U.S. *Digital Millennium Copyright Act* suggests that the notice-and-takedown system used in that context captures large volumes of lawful content.¹⁴³ The UN Special Rapporteur on freedom of expression has also recently referred to aspects of the Network Enforcement Act as “problematically vague” and noted that its failure to define key terms “undermines the claim that its requirements are consistent with international human rights law.”¹⁴⁴ Canadian lawmakers considering law reform initiatives in line with those undertaken in Europe are likely to benefit from a careful study of these efforts and their consequences.

Counterspeech and public education

Civil liberties advocates have long argued that counterspeech offers the most effective and rights-protective measure in response to hate speech, whether on or offline. Rather than putting the government or administrative tribunals in a position to police and censor, they evoke the position once famously expressed by U.S. Supreme Court Justice Louis Brandeis: “if there be time to expose through discussion, the falsehoods and fallacies, to avert the evil by the processes of education, the remedy to be applied is more speech, not enforced silence.”¹⁴⁵ This was part of the position taken by the Canadian Civil Liberties Association at the recent Committee hearings on hate speech online.¹⁴⁶

Critics and courts have been somewhat skeptical of this approach. For example, in *Lemire, Evans J.A.* recently noted on behalf of a unanimous Federal Court of Appeal that “because of the extreme nature of prohibited hate speech it strikes me as fanciful to imagine that those who engage in it are likely to be open to an educative exchange of ideas.”¹⁴⁷ However, advocates of counterspeech are quick to point out that the party communicating hate speech is rarely the intended audience. Instead—and particularly in the context of interventions through public online platforms—the goal is to undermine and denounce the speaker’s views, reveal weaknesses in their position and create social disincentives for those who might otherwise be influenced by hate speech to share it more widely.

Though it has been a longstanding tactic among social movement activists,¹⁴⁸ there is not yet conclusive research regarding whether counterspeech is an effective tool for reducing and responding to hate speech online. This area is nonetheless a rapidly emerging area of study, particularly in the context of online platforms, and has been promoted by the technology industry as an alternative to more costly or restrictive forms of regulation.¹⁴⁹ The research that does exist has generally concluded that variables such as the identity of the counterspeaker, forum, group size, tone, use of humour, perceived credibility, degree of “civility,” and use of imagery can all play a role in determining whether counterspeech is effective in a given context.^{150, 151} There is also an emerging body of research surrounding the use of automated technologies to respond to (rather than to remove) harmful speech online.¹⁵²

In this vein, some have emphasized the need for better reporting tools on social media platforms, which would allow users to more easily flag hate speech. Survey data suggests that a significant majority of internet users in Canada believe that taking these kinds of individual actions is at least somewhat effective approach to countering harmful speech online, and many individuals participate in those activities.¹⁵³ On certain platforms, like Wikipedia and Reddit, users voluntarily play a massive role in content moderation, and researchers are beginning to study how these individuals approach the complex work of balancing civil liberties, community norms and user safety.¹⁵⁴ Most major social media platforms also use extensive (and frequently outsourced) human moderation to manage illegal content, though we are only beginning to understand the psychological consequences of that work on the individuals who perform it.¹⁵⁵

Conclusion

This report can only offer a preliminary overview of the issues. It is by no means exhaustive, and does not aspire to resolve the complex legal and policy questions surrounding hate speech in Canada. However, as the Commission continues to develop its understanding of the problem, this report may provide a useful resource for further study, reflection and public engagement. In the spirit of the Commission's commitment to civil dialogue and stakeholder engagement, a list of discussion questions for various audiences has been provided below.

Discussion questions

For around the dinner table

Have you ever witnessed hate speech in your own life, whether on the internet or somewhere else? What was said? How did it make you feel? Did you, or others, respond? Is there something that you would have done differently if it happened again today?

Is your family, or are certain members of your family, part of a community that has been targeted by hate speech or hate crimes in the past? Is there a story you can share about that experience?

Sometimes there are people in our own families who express hatred and contempt towards another group. It can be difficult to have compassionate conversations with those individuals about their beliefs. What are strategies that we can use to challenge and confront these opinions when they're expressed by someone we care about?

A recent survey indicates that most people in Canada have low levels of trust in any individual actor to regulate and respond to issues like disinformation or hate speech on social media—regardless of whether that actor is a government agency, the social media platform or users themselves.¹⁵⁶ Whose responsibility is it to address issues like hate speech? Who do you trust to balance competing rights, including freedom of expression? How might these institutions become more trustworthy and accountable?

For the high school classroom

In your opinion, is it true that some types of speech are more important to protect in a democratic society than others? What kinds of speech? How do we know? Who should have the right to draw the line?

According to the law in Canada, what is hate speech and what makes it different from other kinds of expression? Take a look at the “hallmarks of hate” that the Human Rights Tribunal describes in *Warman v. Kouba* and share your reflections.

Many civil liberties advocates say that the best way to respond to hate speech is not by giving the government more power to police what individuals can and can't say. Instead, they argue that we should focus on *counterspeech*—which means directly challenging and refuting harmful speech when it happens.

Who has the power to engage in counterspeech, and who doesn't? Whose responsibility is it to “talk back”?

In your experience, is counterspeech an effective way to respond to hateful comments? What effect might it have on the person who made the harmful comments? What might it have on bystanders and victims?

What are some strategies that you use to challenge harmful speech in your own life? Compare your list with [this comic strip](#).¹⁵⁷

In Canada, certain kinds of hate speech are a criminal offence. The consequences of a criminal sentence can be very serious—a conviction can result in prison time and compromise a person's ability to work or travel in the future. Should we ever put someone in prison for something they've said?

For lawmakers and political representatives

Certain groups (and in particular racial, Indigenous, ethnic, linguistic, sexual orientation, gender identity and religious groups) are disproportionately targeted by hate speech. However, these communities are also underserved by law enforcement, and in the case of some groups, subject to systemic discrimination and violence by law enforcement and in prisons. How does this reality complicate greater enforcement of *Criminal Code* prohibitions on hate speech? What should legislatures do in response?

The JUST Committee has recommended the reintroduction of former section 13 of the *CHRA*. Is this the right approach? Before moving forward, what lessons can we learn from the history of legal and political contestation surrounding that provision? What is the problem that Parliament is trying to solve? Can you imagine other ways to respond to that problem?

The JUST Committee has recommended new rules to compel online platforms to identify and remove hate speech. What are the Charter implications—for speech, privacy and other rights—of adopting such measures?

When government makes a decision to regulate technology companies and control speech online, it provokes complex legal debates about everything from freedom of expression and privacy rights to procedural fairness, international human rights law and the limits of jurisdiction. What experts or stakeholders need to be around the table to ensure that any new legislative proposals are constitutional, feasible, necessary and proportionate?

In discussions about regulating online speech, there is often a distinction drawn between “hate speech” and “harmful speech.” While hate speech may attract criminal or administrative sanctions, harmful speech is less well defined, and legislating in this area raises constitutional questions. What do you think are the important differences between hate speech and harmful speech? What role, if any, do you think the government has to play in addressing harmful speech online?

For the community potluck, organizing committee or faith group

When a person promotes hatred, what should justice look like in response? Among the law's many responses to hate speech—from incarceration to fines, injunctions, compensation, settlement and mediation—what kind of process best reflects the values of your community? What kinds of outcomes should the law seek to achieve for victims and offenders? Is there a role for restorative justice in responding to hate speech?

Many groups have called for greater criminal law enforcement and policing of hate speech laws. But historically, laws that restrict speech have often been used to silence vulnerable and marginalized groups, rather than to protect them (for example: the use of anti-pamphleting laws to target religious minorities, the use of mass arrests to silence political protesters, or the criminalization of LGBTQ literature as “obscenity”). What do you make of this tension? Do you trust the government and law enforcement to address issues like hate speech and hate crime in your community?

What is the relationship between words and actions? What is the relationship, in your view, between hate speech and hate incidents or hate crimes? What are effective ways to challenge harmful speech as a community?

For the tech company boardroom

According to survey data, imposing content moderation obligations on social media platforms—including rules that would require the deletion of unlawful content, including hate speech—are generally supported by internet users in Canada.¹⁵⁸ In your view, what is your company's responsibility to flag, filter or remove this kind of content? If this kind of obligation were imposed, what measures (legal, technical or in terms of organizational culture) would need to be put in place to ensure that your company wasn't accidentally removing lawful speech?

Some groups have called for new regulations that would compel online platforms to identify and remove hate speech. What are the ethical and human rights implications of building these capabilities into online platforms? How could these capabilities be abused by a government or a company like yours?

Endnotes

- 1 For a fascinating account of the way language has been used to discuss and respond to hate crime in Canada, see Allyson M. Lunny. 2017. *Debating Hate Crime: Language, Legislatures, and the Law in Canada*. Vancouver: UBC Press.
- 2 See Chris Tenove, Heidi J. S. Tworek and Fenwick McKelvey. 2018. [Poisoning Democracy: How Canada Can Address Harmful Speech Online](#). Public Policy Forum.
- 3 Susan Benesch. et al. January 2002. [Dangerous Speech: A Practical Guide](#). Dangerous Speech Project.
- 4 *Criminal Code*, RSC 1985, c C-46, s 319(2).
- 5 *Criminal Code*, RSC 1985, c C-46, s 319(7) and s 318(4).
- 6 See also *R v Krymowski*, 2005 SCC 7, paras 17-20.
- 7 *Criminal Code*, RSC 1985, c C-46, s 320(7).
- 8 *Mugesera v Canada*, 2005 SCC 40, para 102.
- 9 *R v Keegstra*, [1990] 3 SCR 697, 776, 1 CR (4th) 129.
- 10 *R v Harding*, [2001] OJ No 4953, paras 44-66, 207 DLR (4th) 686.
- 11 *R v Ahenakew*, 2008 SKCA 4.
- 12 *Mugesera*, para 105.
- 13 *Criminal Code*, RSC 1985, c C-46, s 319(3).
- 14 See discussion in *Keegstra* at p 780.
- 15 Genocide is defined in section 318(2) of the *Criminal Code* as: “any of the following acts committed with intent to destroy in whole or in part any identifiable group, namely, (a) killing members of the group; or (b) deliberately inflicting on the group conditions of life calculated to bring about its physical destruction.”
- 16 See also *Mugesera*, para 147.
- 17 *Criminal Code*, RSC 1985, c C-46, ss 318, 430(4.1).
- 18 *Criminal Code*, RSC 1985, c C-46, s 718.2(a) (i).
- 19 See also *R v Presseault*, 2007 QCCQ 384 at paras 47 et seq.
- 20 Statistics Canada, Table 35-10-0066-01. [Police-reported hate crime, by type of motivation, Canada \(selected police services\)](#).
- 21 See also Canadian Race Relations Foundation. March 2020. [Hate Crime in Canada](#).
- 22 House of Commons. June 2019. [Taking Action to End Online Hate: Report of the Standing Committee on Justice and Human Rights](#), p 21. See also footnote 86: “Information provided to the Committee by Statistics Canada by email. Statistics Canada states specifically the following: ‘It is important to note that police-reported data on cyber-related hate crimes are an undercount due to the fact that not all police services have been able to provide Statistics Canada with information on those incidents that are cyber related.’”
- 23 *Ibid.*, 19.
- 24 *Ibid.*, 19-20.

- 25** See *Human Rights Code*, RSBC 1996 c 210, s 7; *Alberta Human Rights Act*, RSAB 2000, c a-25.5, s 3; *The Saskatchewan Human Rights Code, 2018*, SS 2018, c s-24.2, s 14; *Human Rights Act*, SNWT 2002, c 18, s 13.
- 26** *Canada (Human Rights Commission) v Taylor*, [1990] 3 SCR 892 at p 917, 75 DLR (4th) 577.
- 27** See also *Lemire v Canada (Human Rights Commission)*, 2014 FCA 18, paras 89-106.
- 28** *Taylor*, 31-32.
- 29** *Saskatchewan (Human Rights Commission) v Whatcott*, 2013 SCC 11, paras 126-127.
- 30** See generally: *Insurance Corporation of British Columbia v Heerspink*, 1982 CanLII 27 (SCC), [1982] 2 SCR 145; *Winnipeg School Division No. 1 v Craton*, 1985 CanLII 48 (SCC), [1985] 2 SCR 150; *CN v. Canada (Canadian Human Rights Commission)*, 1987 CanLII 109 (SCC), [1987] 1 SCR 1114; *Quebec (Commission des droits de la personne et des droits de la jeunesse) v. Communauté urbaine de Montréal*, 2004 SCC 30; *Tranchemontagne v. Ontario (Director, Disability Support Program)*, 2006 SCC 14 (CanLII), [2006] 1 SCR 513.
- 31** *Canadian Human Rights Act*, RSC 1985, c H-6, s 13 [CHRA], as repealed by *An Act to amend the Canadian Human Rights Act (protecting freedom)*, SC 2013, c 37.
- 32** For more on the history of section 13, see Canada, Canadian Human Rights Commission October 2008. [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. June 2009. [Special Report to Parliament: Freedom of Expression and Freedom from Hate in the Internet Age](#); Ontario Human Rights Commission. January 2009. [Submission to the Canadian Human Rights Commission concerning section 13 of the Canadian Human Rights Act and the regulation of hate speech on the internet by Richard Moon October 2008](#); Library of Parliament, [Hate Speech and Freedom of Expression: Legal Boundaries in Canada](#), Legal and Social Affairs Division at s 4.2.
- 33** For these perspectives, see House of Commons. June 2019. [Taking Action to End Online Hate](#).
- 34** *International Covenant on Civil and Political Rights*, 999 UNTS 171.
- 35** *International Convention on the Elimination of All Forms of Racial Discrimination*, Can. T.S. 1970 No. 28, Art. 4.
- 36** *Canadian Charter of Rights and Freedoms*, s 2(b), Part 1 of *The Constitution Act, 1982*, Schedule B to the *Canada Act 1982 (UK)*, 1982, c 11 [Charter].
- 37** *Canadian Bill of Rights*, SC 1960, c 44, s 1(d) and (f).
- 38** *Charter of Human Rights and Freedoms*, RSQ, c C12, s 3.
- 39** See for example, *Broadcasting Act*, SC 1991, c 11, s 2(3), 35(2) [in public broadcasting]; *Canada Evidence Act*, RSC, 1985, c C-5, s 39.1 [protecting journalistic sources]; *Criminal Code*, RSC 1985, c C-46, s 486.5 [regarding restrictions on publication in criminal proceedings]; *The Saskatchewan Human Rights Code*, 2018, SS 2018, c S-24.2, ss 4, 14(2); *Human Rights Act*, RSY 2002, c 116, s 4.
- 40** *International Covenant on Civil and Political Rights*, 999 UNTS 171, art 19.
- 41** *Universal Declaration of Human Rights*, GA Res 217A (III), UNGAOR 3rd Sess, Supp No 13, UN Doc A/810 (1948) 71 art 19.
- 42** *Charter* at s 2(b).
- 43** For an interesting discussion regarding “freedom of the press” specifically, see dissenting reasons of Abella J. in *R v Vice Media Canada Inc*, 2018 SCC 53, at paras 109 et seq.
- 44** *Charter* at s 2.
- 45** *Thomson Newspapers Co v Canada (Attorney General)*, [1998] 1 SCR 877 at para 81, 159 DLR (4th) 385.
- 46** *Irwin Toy Ltd v Quebec (Attorney General)*, [1989] 1 SCR 927 at pp 967-971, 58 DLR (4th) 577.
- 47** *R v Lucas*, [1998] 1 SCR 439, paras 25-27, 157 DLR (4th) 423; *R v Sharpe*, 2001 SCC 2 at paras 24-27; *R v Barabash*, 2015 SCC 29, paras 15-17; *RWDSU v Dolphin Delivery Ltd*, [1986] 2 SCR 573, 586-587, 33 DLR (4th) 174; *Greater Vancouver Transportation Authority v Canadian Federation of Students—British Columbia Component*, 2009 SCC 31, para 38.

- 48 *Canada (Attorney General) v JTI-Macdonald Corp*, 2007 SCC 30, para 60.
- 49 *R v Zundel*, [1992] 2 SCR 731, para 36, 95 DLR (4th) 202.
- 50 *R v Lucas*, [1998] 1 SCR 439, para 25, 157 DLR (4th) 423.
- 51 *Charter* at s 1.
- 52 See generally *R v Oakes*, [1986] 1 SCR 103, 26 DLR (4th) 200. Note that the Supreme Court has taken a somewhat different approach in the context of administrative decision-making; see *Doré v Barreau du Québec*, 2012 SCC 12, para 6, 37; *Loyola High School v Quebec (Attorney General)*, 2015 SCC 12, paras 39-42; *Trinity Western University v British Columbia College of Teachers*, 2001 SCC 31, para 94; *Whatcott*, para 130; *Canada (Minister of Citizenship and Immigration) v Vavilov*, 2019 SCC 65, para 57.
- 53 See generally *Ford v. Quebec (Attorney General)*, [1988] 2 SCR 712, 54 DLR (4th) 577; *Irwin Toy Ltd v Quebec (Attorney General)*, [1989] 1 SCR 927, 58 DLR (4th) 577 [*Irwin Toy*]; *Thomson Newspapers Co v Canada (Attorney General)*, [1998] 1 SCR 877, 159 DLR (4th) 385.
- 54 *Ford v Quebec (Attorney General)*, [1988] 2 SCR 712, para 56, 54 DLR (4th) 577, quoting Sharpe, R. J. 1987. Commercial Expression and the Charter. 37 *U of TLJ* 229, at 232.
- 55 Compare, for example, *Canada (Attorney General) v JTI-Macdonald Corp*, 2007 SCC 30, paras 70-95 and *Greater Vancouver Transportation Authority v Canadian Federation of Students—British Columbia Component*, 2009 SCC 31.
- 56 *R v Andrews*, 1988 CanLII 200 (ON CA), 65 OR (2d) 161 [*Andrews*].
- 57 *R v Keegstra*, 1990 3 SCR 697.
- 58 *Taylor*.
- 59 *Ibid.*, 915.
- 60 *Keegstra*, 731.
- 61 See also *Irwin Toy*, 970.
- 62 *Keegstra*, 743.
- 63 Special Committee on Hate Propaganda in Canada. 1966. Report of the Special Committee on Hate Propaganda in Canada, 11-15. Ottawa: Queen's Printer. Cited in *Keegstra*, 847.
- 64 *Keegstra*, 746.
- 65 *Taylor*, 918.
- 66 See also *Canada (Human Rights Commission) v Winnicki*, 2005 FC 1493, para 30.
- 67 *Whatcott*, para 74, citing *Mugesera*, para 147.
- 68 Canada, Special Committee on Hate Propaganda in Canada, Report of the Special Committee on Hate Propaganda in Canada, 8. Cited in *Keegstra*, 747.
- 69 This idea is indeed specifically rebutted in *Keegstra*, 763: "Indeed, expression can be used to the detriment of our search for truth; the state should not be the sole arbiter of truth, but neither should we overplay the view that rationality will overcome all falsehoods in the unregulated marketplace of ideas." McLachlin J. responds to this critique in the dissent at p. 803: "While freedom of expression provides no guarantee that the truth will always prevail, it still can be argued that it assists in promoting the truth in ways which would be impossible without the freedom. One need only look to societies where free expression has been curtailed to see the adverse effects both on truth and on human creativity. It is no coincidence that in societies where freedom of expression is severely restricted truth is often replaced by the coerced propagation of ideas that may have little relevance to the problems which the society actually faces. Nor is it a coincidence that industry, economic development and scientific and artistic creativity may stagnate in such societies."
- 70 *Keegstra*, 787.
- 71 *Whatcott*, para 120.
- 72 *Keegstra*, 852, McLachlin J. dissenting.
- 73 *Keegstra*, 859, McLachlin J. dissenting.
- 74 Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. October 9, 2019. Prepared by David Kaye, UNGA, A/74/486, para 1.

- 75 *Whatcott*, para 75.
- 76 *Keegstra*, 853, McLachlin J. dissenting.
- 77 See relatedly, Marc Lemire’s longstanding “attempts to portray himself as a free speech martyr during his lengthy battle with the Canadian Human Rights Commission,” as discussed in Mack Lamoureux. May 9, 2019. [Hamilton Promises Investigation Into Employment of Former Neo-Nazi Leader](#). Vice.
- 78 *Keegstra*, 861, McLachlin J. dissenting.
- 79 *Taylor*, 968.
- 80 *Taylor*, 925, 928, 929.
- 81 *Andrews*, 179.
- 82 See, for example, the “hallmarks of hatred” as summarized in *Warman v Kouba*, 2006 CHRT 50, paras 23-76. These hallmarks include: portraying the targeted group as a powerful menace that is taking control of the major institutions in society and depriving others of their livelihoods, safety, freedom of speech and general well-being; using true stories, news reports, pictures and references from purportedly reputable sources to make negative generalizations about the targeted group; portraying the targeted group as preying upon children, the aged, the vulnerable, etc.; blaming the targeted group for the current problems in society and the world; portraying the targeted group as dangerous or violent by nature; conveying the idea that members of the targeted group are devoid of any redeeming qualities and are innately evil; communicating the idea that nothing but the banishment, segregation or eradication of this group of people will save others from the harm being done by this group; de-humanizing the targeted group through comparisons to and associations with animals, vermin, excrement and other noxious substances; the use of highly inflammatory and derogatory language in the messages to create a tone of extreme hatred and contempt; trivializing or celebrating past persecution or tragedy involving members of the targeted group; and calls to take violent action against the targeted group.
- 83 *Keegstra*, 778, McLachlin J. dissenting.
- 84 *Whatcott*, para 89.
- 85 *Whatcott*, para 95.
- 86 *R v Topham*, 2017 BCSC 259.
- 87 Facebook (November 2019). [Community Standards Enforcement Report](#). Note that Facebook defines hate speech as “violent or dehumanizing speech, statements of inferiority, calls for exclusion or segregation based on protected characteristics, or slurs. These characteristics include race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disability or disease.”
- 88 Google. Transparency Report. [YouTube Community Guidelines enforcement](#). (Accessed March 2020).
- 89 *Ibid*.
- 90 See YouTube. June 2019. [Our ongoing work to tackle hate](#).
- 91 Ryerson Leadership Lab. 2019. [Rebuilding the Public Square](#), 39, 43.
- 92 Commission des droits de la personne et des droits de la jeunesse (Québec). August 2019. Les actes haineux à caractère xénophobe, notamment Islamophobe, Étude présentée dans le cadre du *Plan d'action gouvernemental 2015-2018: La radicalisation au Québec*, p. 108-112.
- 93 See also Maryse Potvin. 2017. [Discours racistes et propagande haineuse. Trois groupes populistes identitaires au Québec](#). *Diversité urbaine*, 17(1), 49-72.
- 94 Becky Gardiner et al. April 12, 2016. [The dark side of Guardian comments](#). The Guardian.
- 95 Lucy Westcott. 2019. [‘The threats follow us home’: Survey details risks for female journalists in U.S., Canada](#). Committee to Protect Journalists.
- 96 *Lemire*, para 9.
- 97 The application of Section 13 to communication over the internet was clarified in the Anti-terrorism Act, S.C. 2001, c. 41 following the decision in *Citron et al. v Zündel*, Canadian Human Rights Tribunal, TD 1/02, 18 January 2002.

- 98 *Lemire*, para 60.
- 99 *Ibid.*, para 61 [emphasis added].
- 100 *Ibid.*, para 68.
- 101 Innovation, Science and Economic Development Canada. June 25, 2019. [Canada's Digital Charter: Trust in a digital world](#).
- 102 *Ibid.*
- 103 New Zealand Foreign Affairs and Trade. [Christchurch Call](#). (Accessed March 17, 2020).
- 104 House of Commons. [Taking Action to End Online Hate](#), 5, citing Centre for Israel and Jewish Affairs, 29 November 2018, Press Release: CIJA Urges Action in Response to Spike in Antisemitic Hate Crimes.
- 105 The Evangelical Fellowship of Canada. February 4, 2019. Calling Parliament to address online hate: Letter to the Minister of Justice.
- 106 See also: House of Commons. [Taking Action to End Online Hate](#), 9-10.
- 107 *Ibid.*, 15-17.
- 108 *Ibid.*, 24-25.
- 109 *Ibid.*, 36-37.
- 110 *Ibid.*, 35-36.
- 111 *Ibid.*, 21-23.
- 112 *Ibid.*, 27.
- 113 *Ibid.*, 29.
- 114 *Ibid.*, 31.
- 115 See conservative party dissenting opinion and statement from Canadian Civil Liberties Association in *ibid.*, 55.
- 116 Innovation, Science and Economic Development Canada. 2020. [Canada's communications future: Time to act](#). Broadcasting and Telecommunications Legislative Review, Ottawa: ISED, rec 95.
- 117 House of Commons. [Taking Action to End Online Hate](#), 39-42.
- 118 See for example, Facebook, [Community Standards: 12. Hate Speech](#) (accessed March 20, 2020); YouTube, [Hate Speech Policy](#) (accessed March 20, 2020); reddit, [Do not post violent content](#) (accessed March 20, 2020); European Commission. June 2016. [The EU Code of conduct on countering illegal hate speech online](#).
- 119 Report of the Special Rapporteur, para 33.
- 120 *Canada (Human Rights Commission) v Liberty Net*, [1998] 1 SCR 626, [1998] SCJ No 31, para 50.
- 121 United Nations Office of the High Commissioner. 2011. [Guiding Principles on Business and Human Rights](#). HR/PUB/11/04., Geneva: United Nations.
- 122 See for example, *EU Code of Conduct*; Evelyn Douek. 2019. [Facebook's 'Oversight Board': Move Fast with Stable Infrastructure and Humility](#). 21 N.C. J. L. & Tech. 1; Shirin Ghaffary. January 28, 2020. [Here's how Facebook plans to make final decisions about controversial content it's taken down](#) Vox; Sarah Perez. May 15, 2018. [Facebook's new transparency report now includes data on takedowns of 'bad' content, including hate speech](#). TechCrunch.
- 123 Paul Mozur. October 15, 2018. [A Genocide Incited on Facebook, With Posts From Myanmar's Military](#). The New York Times.
- 124 Evelyn Douek. October 22, 2018. [Facebook's Role in the Genocide in Myanmar: New Reporting Complicates the Narrative](#). Lawfare Blog.
- 125 See generally Siena Anstis et al. September 2019. [Annotated Bibliography: Dual-Use Technologies: Network Traffic Management and Device Intrusion for Targeted Monitoring](#). Citizen Lab.
- 126 Jonathon Penney et al. 2018. [Advancing Human-Rights-By-Design In The Dual-Use Technology Industry](#). Columbia Journal of International Affairs, 71(2), 103-110.
- 127 Jakub Dalek et al. 2018. [Planet Netsweeper](#). The Citizen Lab.
- 128 Ronald Deibert, Joshua Oliver, and Adam Senft. 2019. [Censors Get Smart: Evidence from Psiphon in Iran](#). Rev Policy Research, 36(3), 341-336.

- 129** Lotus Ruan, Jeffrey Knockel, and Masashi Crete-Nishihata. March 3, 2020. [Censored Contagion: How Information on the Coronavirus is Managed on Chinese Social Media](#). The Citizen Lab.
- 130** Note that chilling effects of perceived surveillance on freedom of expression has been demonstrated in other contexts. See for instance Jonathon Penney. 2017. [Internet surveillance, regulation, and chilling effects online: a comparative case study](#). *Internet Policy Review*, 6(2).
- 131** See also the potential benefits of anti-cyberbullying legislation on women’s expression: Danielle Keats Citron and Jon Penney. 2019. [When Law Frees Us to Speak](#). *Fordham Law Review*, 87(6), 2317.
- 132** Amnesty International. 2017. [Protecting Human Rights on the Internet](#).
- 133** Nathalie Maréchal and Ellery Roberts Biddle. 2020. [It’s Not Just the Content, It’s the Business Model: Democracy’s Online Speech Challenge](#). Ranking Digital Rights.
- 134** See Maarten Sap et. al. 2019. [The Risk of Racial Bias in Hate Speech Detection](#). Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 1668-1678; Thomas Davidson, Debasmita Bhattacharya, and Igmarr Weber. 2019. [Racial Bias in Hate Speech and Abusive Language Detection Datasets](#). Proceedings of the Third Workshop on Abusive Language Online, Association for Computational Linguistics; Anti-Defamation League. 2018. [The Online Hate Index](#). Centre for Technology and Society; Thomas Davidson et al. 2017. Automated Hate Speech Detection and the Problem of Offensive Language. ICWSM 2017.
- 135** Stefanie Ullmann and Marcus Tomalin. 2019. [Quarantining online hate speech: technical and ethical perspectives](#). *Ethics and Information Technology*, 22, 69-80.
- 136** Angela Chen. September 26, 2019. [Limiting message forwarding on WhatsApp helped slow disinformation](#). MIT Technology Review.
- 137** Report of the Special Rapporteur, para 51.
- 138** Some of these arguments were raised in a different context (albeit largely unsuccessfully), by some interveners in *Google Inc v Equustek Solutions Inc*, 2017 SCC 34. See factums on appeal [online](#).
- 139** A few comparative examples: Chinmayi Arun & Nakul Nayak. December 2016. [Preliminary Findings on Online Hate Speech and the Law in India](#). Berkman Klein Research Publication No 2016-19; Niousha Roshani. December 2016. [Grassroots Perspectives on Hate Speech, Race, & Inequality in Brazil & Colombia](#). Berkman Klein Research Publication No 2016-18; Gayathry Venkiteswaran. October 2017. [“Let the Mob Do the Job”: How Proponents of Hatred are Threatening Freedom of Expression and Religion Online in Asia](#). Association for Progressive Communications; Inter-American Commission on Human Rights. 2015. [Hate speech and incitement to violence against lesbian, gay, bisexual, trans and intersex persons in the Americas](#).
- 140** Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (“Network Enforcement Act” “NetzDG”). June 30, 2017. Deutscher Bundesrat: Drucksachen (Br-Drs) 536/17.
- 141** France, Proposition de loi visant à lutter contre la haine sur internet, n° 1785, déposé le 20 mars 2019.
- 142** See discussion in Chris Tenove, Heidi J. S. Tworek and Fenwick McKelvey, *Poisoning Democracy* at s 3.1.
- 143** See Jennifer M. Urban, Joe Karaganis and Brianna L. Schofield. 2017. [Notice and Takedown in Everyday Practice](#). UC Berkeley Public Law Research Paper No. 2755628.
- 144** Report of the Special Rapporteur, para 33.
- 145** *Whitney v California*, 274 US 357 (1927).
- 146** See quote from Cara Zwibel in [Taking Action to End Online Hate](#), 55.
- 147** *Lemire*, para 65.

- 148** See for example, Engy Abdelkader. 2014. *Savagery in the Subways: Anti-Muslim Ads, the First Amendment, and the Efficacy of Counterspeech*. *Asian American Law Journal*, 21(43); Rachel Briggs and Sebastien Feve. 2013. *Review of programs to counter narratives of violent extremism*. Institute of Strategic Dialogue; Robert D. Richards and Clay Calvert. 2000. *Counterspeech 2000: A New Look at the Old Remedy for “Bad” Speech*. *BYU Law Review*.
- 149** See, for example, comments by Facebook COO Sheryl Sandberg in Kurt Wagner. January 21, 2016. [Want to Combat Hate Speech on Facebook? Try a ‘Like Attack.’ Says COO Sheryl Sandberg](#). Vox.
- 150** For more information on online counterspeech, see Cathy Buerger’s excellent literature review on the subject: Cathy Beuger. February 20, 2020. [Counterspeech: A Literature Review](#). Dangerous Speech Project.
- 151** See also Daniel Jones and Susan Benesch. August 9, 2019. [Combating Hate Speech Through Counterspeech](#). Berkman Klein.
- 152** For example, one researcher has found that on Twitter, white men using racist slurs who were confronted by a bot that appeared to be a white counterspeaker with a large number of followers were more likely to change their behaviour than when confronted by a bot that appeared to be a black person, or that had few followers; Kevin Munger. 2017. *Tweetment effects on the tweeted: Experimentally reducing racist harassment*. *Political Behavior*, 39(3), 629-649.
- 153** *Rebuilding the Public Square*, 42.
- 154** See for example Robyn Caplan. 2018. [Content or Context Moderation? Artisanal, Community-Reliant, and Industrial Approaches](#). *Data & Society*; J. Nathan Matias. 2019. [The Civic Labor of Volunteer Moderators Online](#). *Social Media + Society*, 5(2); Justin Clark et al. 2019. [Content and Conduct: How English Wikipedia Moderates Harmful Speech](#). Berkman Klein Research Publication No 2019-5; Björn Ross et al. 2016. [Measuring the Reliability of Hate Speech Annotations: The Case of the European Refugee Crisis](#). *Proceedings of NLP4CMC III: 3rd Workshop on Natural Language Processing for Computer-Mediated Communication (Bochum)*, Bochumer Linguistische Arbeitsberichte, 6-9.
- 155** See generally: Sarah T. Roberts. 2019. *Behind the Screen: Content Moderation in the Shadows of Social Media*. Yale University Press; *The Cleaners*. 2018. Directed by Moritz Riesebeck and Hans Block. Gebrueder Beetz Filmproduktion; Casey Newton. January 24, 2020. [YouTube moderators are being forced to sign a statement acknowledging the job can give them PTSD](#). The Verge.
- 156** *Rebuilding the Public Square*, 52.
- 157** See also Rachel Brown, 2016. [Defusing Hate: A Strategic Communication Guide to Counteract Dangerous Speech](#); Susan Benesch et al. October 14, 2016. [Considerations for Successful Counterspeech](#). Dangerous Speech Project.
- 158** *Rebuilding the Public Square*, 56.

