

#BLOCKHATE



Centering Survivors and Taking Action
on Gendered Online Hate in Canada -
National Report

OCTOBER 2022



Land Acknowledgment

The author and project team acknowledge the privilege and responsibility of living and working on the traditional territory of Indigenous Nations. YWCA Canada's national office is located in Tkaronto, also known as Toronto, the traditional territory of the Wendat, the Anishinaabeg, Haudenosaunee and Métis. We acknowledge the enduring presence of Indigenous peoples in this area and recognize the current treaty holders, the Mississaugas of the New Credit First Nation, as contemporary stewards of the land.

The harm that settler-colonialism has perpetrated and continues to perpetrate on First Nations, Inuit, and Métis people extends beyond physical environments. Digital platforms are among the current sites of ongoing colonial discrimination and dispossession. As a legacy institution that benefited from and contributed to colonial policy, YWCA Canada is committed to centering decolonial efforts and supporting the work of Indigenous communities, Elders, and knowledge keepers. Through our work on gender equity, our objective is to advance the safety, worldviews, and self-determination of Indigenous people in online and offline spaces.

About YWCA Canada

YWCA Canada is a leading voice for women, girls, Two-Spirit and gender-diverse people. For over 150 years, we've been at the forefront of a movement: to fight gender-based violence, build affordable housing and advocate for workplace equity. We work to advance gender equity by responding to urgent needs in communities, through national advocacy and grassroots initiatives. Today, we engage young leaders, diverse communities, and corporate partners to achieve our vision of a safe and equitable Canada for all.

Acknowledgments

We extend deep appreciation to the survivors and young women and gender diverse youth that participated in focus group consultations and surveys and trusted us with their truth. This report was made possible by their contributions. Our gratitude goes to the eight YWCA Canada Disrupt and Dismantle! Leadership Fellows whose active engagement and shared commitment steered the research process. The project benefited significantly from the insights and cross-sectoral experiences of the Block Hate advisory committee members. We are grateful to the following reviewers for generously providing substantive feedback throughout the development of this report (in alphabetical order by last name): Alejandro Mayoral Baños, Cathy Buerger, Nadia Duman, Nammal Khan and Timothy Quinn. YWCA Canada gratefully acknowledges Anishinaabe Onyota'a:ka artist Tsista Kennedy and the Indigenous Friends Association for the artwork featured in this report.

The Block Hate project is funded by the generous support of Public Safety Canada's Community Resilience Fund. Several organizations partnered with YWCA Canada on different facets of the research project. The Canadian Race Relations Foundation (CRRF) and YWCA Canada launched the #BlockHate campaign to raise awareness around the rise of hate speech and racism online. National surveys were conducted by Abacus Data and Environics Research. This report was researched and written by Jolin Joseph with contributions from Amanda Arella, Gustavo Freire, Reese Clapp and Nammal Khan. The report was coordinated by Raine Liliefeldt, Hamzia Bawa-Zeba, Amanda Arella, Aniska Ali, and Natasa Boskovic.

Executive Lead:

Raine Liliefeldt

Director of Member Services & Development, Interim CEO, YWCA Canada

Project Coordination:

Aniska Ali

Director of Philanthropy and Strategic Initiatives, YWCA Canada

Hamzia Bawa-Zeba

National Project Coordinator, YWCA Canada

Research Lead and Report Author:

Jolin Joseph

Manager of Research, Education & Knowledge Mobilisation, YWCA Canada

Report Contributor:

Amanda Arella

Director of Public Policy, Advocacy & Strategic Communications, YWCA Canada

Additional Contributors:

Gustavo Freire, Reese Clapp, Nammal Khan

Acknowledgments

Design:

Artwork by Tsista Kennedy, Anishinaabe Onyota'a:ka artist and the Indigenous Friends Association

Report Covers by Feisty Creative

Layout by Amanda Arella, Director of Advocacy, Public Policy & Strategic Communications, YWCA Canada

Translation Support:

A Propos Services

Jenny Cloutier

Bilingual Project Officer, YWCA Canada

Suggested Citation:

Joseph, J. (2022). Centering Survivors and Taking Action on Gendered Online Hate in Canada: National Report. YWCA Canada.

External Reviewers:

Nadia Duman – Work Transition Coordinator, YWCA St. John's and Leadership Fellow

Timothy Quinn – Founder, [Dark Data Project](#)

Advisory Committee Members:

Cathy Buerger – Director of Research, [Dangerous Speech Project](#)

Alejandro Mayoral Baños – Executive Director and Founder, [Indigenous Friends Association](#)

Rachel Curran – Public Policy Manager, [Meta Canada](#)

Amira Elghawaby – Director of Programs and Outreach, [Canadian Race Relations Foundation](#)

Jad El Tal – Director of Research and Policy, [Canadian Arab Institute](#)

Tamar Huggins Grant – CEO and Founder, [Tech Spark](#)

Derek Ruths – Associate Professor of Computer Science, [McGill University](#)

Naoufel Testaouni – CEO, [QueerTech](#)

Denise Williams – CEO, [First Nations Technology Council](#)

Block Hate Disrupt & Dismantle Leadership Fellows:

Jastina Aujla

Érika Cantu

Nadia Duman

Noor Fadel

Houmou Guiro

Pauline Itheme

Racheal Walser

Kelly Weiling Zou

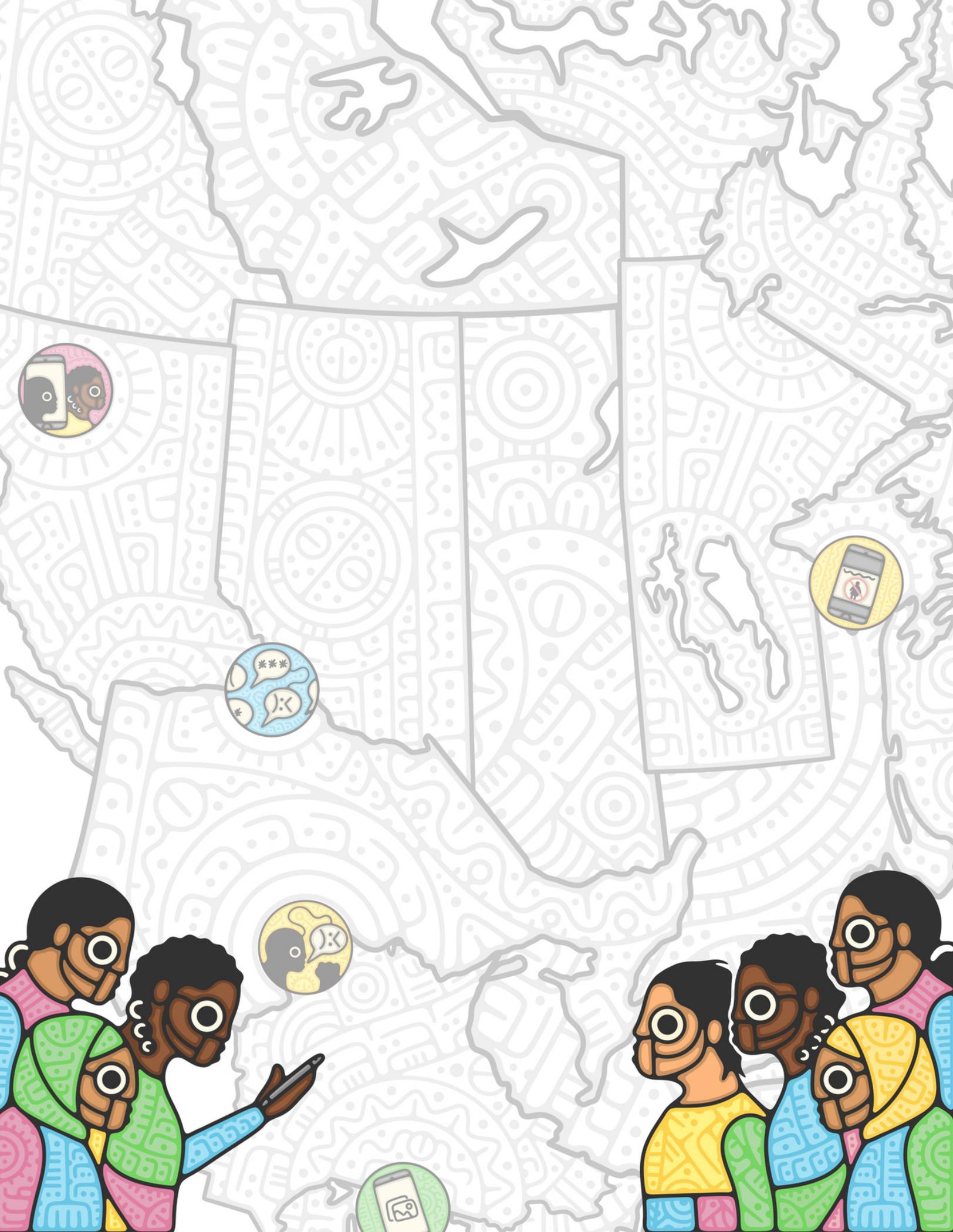


Table of Contents



01

Executive Summary

03

Realities of Online Hate in Canada

Experienced Early and Everywhere

Role of Online Platforms

Whose Freedom?

Gendered Nature of Online Hate in Canada

Deactivating Democracy: The Chilling Effects of Online Hate Circulation

How Youth are Responding

04

Recommendations from Youth Consultations across Canada

Advocacy and Awareness

Accountability

Support

Regulatory Responses to Online Hate

02

Introduction

What YWCA Canada is Doing

Methodology

05

Feminist Regulatory Responses to Online Hate – A Federal Responsibility

Legal Framework

Platform Regulation and Content Moderation

Oversight, Transparency and Accountability

Intersectional Approaches to Addressing Online Hate

Research, Education, and Support



01 Executive Summary

“The Internet is now a common good we all consume. The more we disengage from it, the more it has to run itself. The more involved people can be, the more we can steer the wheel towards certain outcomes.”

- Focus Group Participant and Survivor of Online Hate

The Government of Canada has made a commitment to introduce a regulatory framework for online safety. Access to and safety of online environments are key dimensions of gender equity. Through research with young women and gender diverse youth across Canada, YWCA Canada has outlined critical proposals to take into consideration during the development of legal and regulatory regimes around online hate and related harms. These efforts are necessary to ensure free expression and full participation of young women and gender diverse youth in online platforms and public life.

This report centers the voices, lived experiences, and strategies shared by young women and gender diverse youth aged 16-30 years. YWCA Canada commissioned national surveys and convened ten focus groups to discuss individual and collective experiences of online hate across Canada and develop community-generated, survivor-centric solutions to curb the circulation of digital hate and mitigate its harms.

Emerging from community-based dialogue and data collection, the recommendations included in this report provide ethical guidelines and equitable benchmarks for systems-level changes (including federal action, platform liability, public education, and community responses) to ensure online safety across Canada.

Calls for Action from Youth Consultations

Participants called for a whole systems approach, involving coordinated collective action by multiple sectors and stakeholders to effectively respond to digital hate. These can broadly be categorized under four themes:

1. Advocacy and Awareness
2. Accountability
3. Support
4. Regulatory Responses to Online Hate



Feminist Regulatory Responses to Online Hate – A Federal Responsibility

This report provides fundamental considerations and specific recommendations for safer online spaces emerging from research and community consultations. It offers five key principles that all levels of government must consider as they develop and implement policy and legislative responses to online hate. Building from that, it proposes fifteen recommendations for federal legislative action and commitment organized around five pillars.

Five fundamental considerations:

1. The Government's Role in Online Safety
2. Intersectional and Survivor-Centric Approach
3. Recognition of Individual and Systemic Harm
4. Guarantee Right to Privacy and Participation
5. Definitional Clarity and Legislative Scope

Recommendations for federal action:

Legislative Action

- Define which online platforms and services are covered by regulation and create categories based on the nature of content and potential harms
- Proportional and specific obligations for different categories of online platforms
- Include the regulation of hate speech in addition to hate crimes, address harmful content in addition to illegal content, and respond to the evolving nature of hate online

Platform Regulation and Content Moderation

- Regulatory coverage of public-facing features and private messaging functions of online platforms
- Establish a set of requirements, including proactive and preventative measures; mandatory sensitivity training; accessible and expedient reporting and redressal mechanisms; clear content review and moderation processes; reporting of incidents and resolution
- Ensure content moderation and data sharing are conducted from a survivor-centric, trauma-informed and harm mitigation approach

Oversight, Transparency and Accountability

- Create a centralized Digital Safety regulator to protect and promote online safety
- The Digital Safety regulator should evaluate platform policies and enforce mechanisms to guarantee online social platforms comply with their obligations
- Oblige online platforms to collect and report data on online hate; maintain transparency in how content is reviewed; make timely reports back to individuals who reported or flagged content; undergo independent audits

Intersectional Approaches to Addressing Online Hate

- Recognize online hate as real harm and acknowledge the connections between online hate and offline violence.
- Multi-stakeholder approaches to address the root causes of online hate
- Apply a Gender-Based Analysis (GBA+) lens to understand the intersectional impacts of legislation and policies

Research, Education, and Support

- Advance research, training, capacity-building, community-based programming, public education, and peer-supports through sustained funding
- The mandate of the Digital Safety regulator should include prevention, public education, and democratic engagement
- Resource community organizations to offer context-specific supports, and youth-specific interventions



O2 Introduction

Online social platforms[1] are essential to how we exchange ideas, engage with others, and express and come to understand ourselves. The design and governance of these platforms shape participation in online spaces and public life. They create new avenues for communication, community-building, and commercial activity to flourish. However, because they offer few constraints, online social platforms facilitate the spread of online hate that can foreclose opportunities and undermine the safety and dignity of equity-deserving groups.[2]

Online hate is understood and experienced differently by everyone it impacts. This was evident in the ways young women and gender diverse youth across Canada described online hate. In focus groups conducted by YWCA Canada, youth participants defined online hate speech as “the use of information and communications technology (ICT)” to spread “any sort of rhetoric or comment that targets and vilifies specific identities and individuals”, as a means “to embarrass, to exclude, to threaten, to shame, [and] to humiliate”. The notion that online hate is closely linked to systems of oppression featured prominently across focus group and survey responses conducted as part of YWCA Canada’s Block Hate: Building Resilience against Online Hate project funded by Public Safety Canada’s Community Resilience Fund. Participants indicated additional complexities of hate speech and harassment in the digital realm, including the lack of responsibility and repercussions due to anonymity and insufficient regulatory action.

Hate can be expressed online in various forms—malicious communications (emails, messages and comments that cause distress and anxiety, offensive and inflammatory memes, emojis, symbols, hashtags or jokes, and threats), cyberbullying, cyberstalking, harassment, stoking hatred through content (text, image, video, audio) and incitement to commit violence or harm (to self or others). Hate is also more easily spread online. Online social platforms provide efficient ways to share and transmit attitudes, beliefs, and agendas. Participants spoke of the networked and organized nature[3] of online hate groups that constantly adjust their behaviours and develop coded language[4] to be effective or relevant and avoid detection. Most participants converged on the idea that online platforms are complicit in amplifying and perpetuating online hate through algorithm-driven curation and circulation practices that prioritize specific kinds of content and privilege certain forms of engagement.

Since the onset of the COVID-19 pandemic, online connectedness and collective reliance on platforms for education, business, and personal use have reached unprecedented levels. Schools, colleges, universities, and workplaces around the world embraced the digital shift in operations.[5] As young people spend more time on screens to combat social isolation, they are more susceptible to encountering online hate. [6]

Youth are key constituents on online spaces. As such, they can be targets, instigators, amplifiers and mitigators of online hate. In a 2022 national survey conducted by Environics Research for YWCA Canada, 83 percent of young women and gender diverse youth aged 16-30 considered hate speech to be a problem. [7] A 2019 study by MediaSmarts revealed that Canadian children and youth are increasingly exposed to casual prejudice or 'cultures of hatred' in online environments where racism, sexism, misogyny and other forms of discrimination and prejudice are normalized.[8] Recurring exposure to online hostility and harassment leads young people to becoming less sensitive to online hate and less empathetic to those that are targeted by hateful speech and slurs in the comments on Instagram posts, demeaning videos on TikTok, or Tweets used to negatively target certain marginalized groups.[9]

In fact, research suggests the more often youth witness online acts of hatred, the more likely they are to engage in hate-motivated online behaviours and speech themselves.[10] A 2020 study identified a 70 percent uptick in instances of hate speech, abusive language and cyberbullying, among kids and teens.[11] Focus group participants flagged how easy it is for young people to get caught up in online feedback and validation loops that cause them to adopt hateful, discriminatory, even extremist attitudes. When youth and other platform users are not empowered to recognize and interrupt hate and prejudice online and when they do not see effective and timely platform responses, it contributes to the sense that hateful sentiments are common and acceptable parts of the social norms of the community.[12]

Hateful beliefs and behaviours are demonstrably increasing on online social platforms. A 2016 analysis of online behaviour across Canada by CBC's Marketplace suggests a 600 percent annual increase in the frequency of racist, Islamophobic, sexist or otherwise intolerant language online.[13] Approximately 60 percent of young women and gender diverse people in Canada who directly experienced online hate say they are targeted monthly, if not more frequently and for many survivors it is a daily reality.[14]

Online toxicity is rooted in established systems of social and structural discrimination. Its impacts are intersectional and unevenly experienced. Members of Indigenous, racialized, 2SLGBTQIA+[15] and other equity-deserving groups are disproportionately targeted. A public poll conducted by the Canadian Race Relations Foundation and Abacus Data released in January 2021 revealed that people from racialized communities in Canada are three times more likely to be the target of online expressions of hate and racism than their non-racialized counterparts and that the age group most at risk is young people between 18 to 29.[16] Black, Indigenous, racialized and immigrant youth survivors of online hate involved in focus groups conducted as part of this study said their posts and message inboxes are often filled with a 'cocktail of racism and misogyny'.[17] Further, focus group consultations drew attention to the ways that Anti-Black racism, Anti-Indigenous racism, Islamophobia and xenophobic or other racist discourses and discrimination are baked into the interface, design features, and business models of online social platforms.[18]

Online social platforms are fertile ground for the rise of race-based harassment and violence[19] and technology-facilitated gender-based violence (TFGBV).[20] Drawing on their deep expertise in the gendered impacts of online hate and technology-facilitated gender-based violence, abuse, and harassment, a seminal 2021 report by Women's Legal Education and Action Fund (LEAF) details how platform policies and practices can be a catalyst for violent and prejudicial language, action or ideation against women and gender diverse people and racialized communities.

YWCA Canada posits that online communications, discourses and behaviours that (i) (re)produce and extend unequal gender power relations, (ii) pose barriers for women and gender diverse people's access to and participation in online spaces and public life, and (iii) are directed to individuals and groups on the basis of their real or perceived gender, gender expression, gender identity, often in combination with other related and intersectional marginalized identities such as sexual orientation and racialization, constitutes gendered online hate.[21] Taken together, the gendered and racialized nature of online hate causes incalculable harm and devastating impacts on generations of women and gender diverse youth and Canadian community in general.

Hate disseminated online has far-reaching impacts on those that experience and witness it and broad implications for society. Harms resulting from exposure to online hate are not always obvious.[22] Focus group participants and survey respondents shared that the prevalence and normalization of online hate subjects youth to psychological and physical harm, including increasing vulnerability to anxiety, depression, and isolation from

family, peers and communities they're a part of. The trauma-induced through online hate causes youth from historically and systematically marginalized and equity-deserving groups to withdraw from online platforms and public life. Online humiliation, degradation and targeting, undermines the safety and participation of women, gender diverse people and equity-deserving groups and sends the message that they do not belong on online communities. As a result, they may be more reluctant to take on profile-raising, public facing opportunities.

The real and present links between online hate and offline violence have been discussed and demonstrated.[23] Online hate can be a precursor or predictor of offline harm.[24] Online hate propaganda has been used to signal and organize ethno-racial, religious and sexist or misogynist violence.[25] Offline events such as Canadian federal, provincial, or municipal elections, recent trucker convoys[26], and the murder of a Muslim family in London, Ontario[27] have correlated with and led to increases in hateful activity and radicalization across online platforms.[28] Focus group participants mentioned other highly politicized issues such as the Wet'suwet'en movement, Palestinian resistance, the murder of George Floyd, anti-trans legislation and abortion bans as creating the perfect storm for rising online hate. Studies have shown that online Anti-Black racism both spurred and emerged from the Black Lives Matter movement.[29] The COVID-19 pandemic and related media coverage has been linked to the staggering surge in Anti-Asian hate online and offline.[30]

A 2021 social media data study commissioned by youth charity Ditch the Label reports an increase of Anti-Asian online hate speech by 2,770 per cent in the first year of the pandemic alone.[31] Today, no clear distinction can be made between harms experienced on the streets and on our screens; each have tangible and lasting impacts on individuals and society.

Despite its serious implications, online hate is slippery topic and consequently, less subject to control. In focus groups and survey responses, youth across Canada called for adequate and appropriate survivor-centric regulatory frameworks, counter efforts, community and platform responses to minimize hate and mitigate harms. The objective of YWCA Canada's Block Hate project is to build digital resilience that both includes and goes beyond individual and community resilience or online safety, to systems-level resilience that incorporates (i) re-imagining and regulating the design, governance and content moderation practices of online platforms and (ii) informs the development and implementation of civil society responses and regulatory frameworks that support and enforce safe and inclusive online platform engagement. Emerging from community-based dialogue and data collection, the recommendations included in this report provide ethical guidelines and equitable benchmarks for systems-level changes (including federal action, platform liability, public education, and community responses) to ensure online safety across Canada.



#BLOCK HATE

WHAT THE YWCA IS DOING

Block Hate: Building Resilience against Online Hate Speech is a four-year research and knowledge mobilization project funded by Public Safety Canada's Community Resilience Fund. It works with a variety of sectors to co-create concrete solutions for online hate speech and hate crimes in communities across Canada. Using a participatory community-based research approach, the overall objective of this project is to improve community resilience and develop tools to prevent, address and report online hate across Canada.

It aims to strengthen civil society responses at the national and local levels through engagement with the information, communication, and technology (ICT) sector as well as local communities in creating and sharing counter-narratives against Anti-Black racism, Anti-Indigenous racism, Islamophobia, transphobic, homophobic, xenophobic, radical nationalism, or other harmful discourses. The project focuses on youth aged 14-30 in communities across Canada, particularly youth experiencing marginalization including young women and gender diverse youth, Black, Indigenous, and racialized youth, 2SLGBTQIA+ youth, youth from rural or remote communities and religious minorities, and youth with disabilities.

The project is comprised of four components: research and knowledge mobilization, creating partnerships with a coalition of stakeholders, building tools to combat hate speech in Canada, and capacity building for community actors. Findings will inform conversations with experts and key stakeholders—community groups and organizations, survivors, technology companies, policy makers, academics—to address hate speech, hate crime, and radicalization to violence. The tools developed through this project will help strengthen communities, establish safer online spaces, and reduce the likelihood of violence and radicalization by creating safer, more inclusive online spaces.

The project builds on previous research YWCA Canada conducted around technology-facilitated gender-based violence. As part of Project Shift- Creating a Safer Digital World for Young women, YWCA Canada and project partners developed a range of resources to support young women, girls, and gender diverse youth to stay safer online and for adults to better support young people when an incident of cyberviolence occurs.

Findings and further steps identified by Project Shift informed the intersectional approach to online hate and platform systems and governance taken up in the Block Hate Project.

To support and extend the work of the Block Hate project, YWCA Canada developed the YWCA Disrupt & Dismantle! Leadership Fellowship. Through this paid fellowship, eight young QTBIPOC [32] employees from across the YWCA movement were given opportunities to advance their progressive leadership and engage in participatory data collection.

Working collaboratively on community dialogue and data collection, Leadership Fellows ensured that knowledge gaps and lived experiences were named and centered in the work of YWCA Canada. Fellows participated in pilot focus groups where they shaped and co-created the themes and questions of the research. With training in community-based participatory action research (CBPAR) methods, Fellows led focus groups across the country to understand the unique contexts and experiences of survivors in Canada and co-create community-led solutions to confronting online hate.



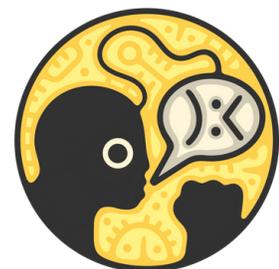
METHODOLOGY

With a focus on community-led learnings and survivor-centric solutions, the Block Hate project partnered with YWCA Disrupt & Dismantle! Leadership Fellows—a cohort of self-identified young, Two-Spirit, queer, trans, Black, Indigenous, and people of color members of the YWCA movement—to take this work forward. The research began with a review of existing online hate research and community and legislative interventions to gather preliminary insights about the state of online hate in Canada and the roles, responsibilities and regulation of online social platforms. This context was followed by two complementary phases: (1) Focus Group Consultations, and (2) National Surveys. Leadership Fellows were included in all facets of the research project and provided training, support and mentorship in community-based participatory action research (CBPAR) methods, anti-racism and anti-oppression. Through their work, Fellows ensured that knowledge gaps, lived experiences, and local circumstances were named and centered in the development of this report.

YWCA Canada convened ten focus groups consultations in February and March 2022, engaging over 50 participants [33] in communities across Canada. Focus groups were conducted in both French and English. Informed consent was obtained from all individual participants involved in this study.

Central to the community-based participatory action research framework is the need to begin with survivors' concerns and proposed action strategies, and to support this with research that puts the insights, issues, and ideas of the communities consulted into calls for action. As youth and survivors themselves, Fellows mapped the unique contexts and experiences of survivors in Canada to co-create community-led solutions that address the rise of online hate.

Adding representative data to these rich peer exchanges, YWCA Canada commissioned a national survey with over 1000 young women and gender diverse youth aged 16-30 conducted by Environics Research between June-September 2022. [34] The data was weighted according to the national census to ensure that the sample matched Canada's population according to age and region. Young women and gender diverse youth across the country were invited to share their understanding, perspectives, experiences, and responses to online hate through the two surveys. The content and recommendations presented in this report center the voices, lived experiences, and potential solutions shared by young women and gender diverse youth aged 16-30 years from across Canada. Their contributions remain the cornerstone of these efforts.





03 Realities of Online Hate Speech in Canada

From focus groups consultations and survey responses, it is clear that the current architecture of online spaces enable and embolden online expressions of hate speech. “[Online platforms] have provided unprecedented opportunities for individuals and groups to broadcast harmful content and recruit adherents to hateful ideologies”. [34A] Participants shed light on how youth experience and respond to unrelenting online hate across Canada.

EXPERIENCED EARLY, AND EVERYWHERE

Across Canada more than 1 in 4 young women and gender diverse youth have been personally targeted by online hate and 1 in every 2 have witnessed it, according to a national survey conducted for YWCA Canada by Environics Research. [35] Approximately 60 percent of young survivors who have directly experienced online hate are targeted monthly, if not more frequently, and for many it is a daily reality. Youth with disabilities are 70 percent more likely to directly experience online hate, youth from 2SLGBTQIA+ and Indigenous communities are about 60 percent more likely, and young Black people are 53 percent more likely to be made targets by online hate.

Young people frequently experience online spaces as toxic or threatening. Half of those who reported experiencing online hate speech in the last couple of years were targeted on Facebook; other common avenues include Instagram, TikTok and Snapchat. [36]

For one focus group participant, her first brush with online hate began as a young child on Club Penguin [37] and continues on Twitch and Twitter.

Across Canada more than **1 in 4** young women and gender diverse youth have been personally targeted by online hate and **1 in every 2** have witnessed it.

Another survivor and avid gamer recounted the time a man attacked her online by releasing where and when she would be at a convention, resulting in her receiving multiple threats of sexual and physical violence. She later found out that a lot of the messages had been sent by teens and pre-teens. According to her, hate proliferating on online platforms “starts young with Twitch and it starts very aggressive.” Eventually she left the platform altogether. This impacted her income and also stopped her from taking up other paid in-person opportunities out of fear that she would be harmed. These experiences reveal that “online hate speech always has an underlying threat of moving offline and threatening to dox and assault” platform users in person. Participants described the trauma and recurring fear associated with online hate. This impacted her income and also stopped her from taking up other paid in-person opportunities out of fear that she would be harmed. These experiences reveal that “online hate speech always has an underlying threat of moving offline and threatening to dox and assault” platform users in person. Participants described the trauma and recurring fear associated with online hate.

As children and youth spend more time on mobile devices, tablets, and computers for academic and non-academic pursuits[38], a 37 percent increase in the overall online victimization of youth was reported by the Canadian Centre for Child Protection in February 2022.[39] As one participant put it, when kids are first exposed to hateful messaging, “they learn without context or meaning.”

A 2020 study revealed a worrying **70 percent rise** in online toxicity, cyberbullying and interpersonal hate speech among children during online chats.

With continued exposure to cultures of hatred[40], they become targets, perpetrators, and amplifiers of online hate. A 2020 study revealed a worrying 70 percent rise in online toxicity, cyberbullying and interpersonal hate speech among children during online chats.[41] While young people in Canada remain at heightened risk of experiencing online hate and harassment, they are also more susceptible to being influenced by hate-based views.

ROLE OF ONLINE SOCIAL PLATFORMS

The inner workings of online platforms influence what users share and what they come to believe. Meta and other for-profit companies engineer proprietary algorithms that implicitly or explicitly stoke divisiveness and hate. Engagement-driven algorithms boost the content most likely to draw clicks, comments, and shares. These patterns are far from passive. Posts surfaced on users' news feed set the stage, or tone, for their interactions and the information they process. For instance, one focus group participant experimented with creating a secondary TikTok account. After she made a single comment about right-wing nationalism, similar content showed up as suggestions. This confirmation bias was identified by other focus group participants who noticed their feeds filling up with perspectives that supported their views and did not challenge or expand them. According to Fisher and Taub, young people follow the 'incentive structures' and social norms of online platforms until they gradually "arrive at hate speech."^[42]

Another insidious aspect of current platform operations is the impact of algorithmic bias^[43] on diversity. Participants noted that platforms selectively amplify content by mainstream content-creators to the detriment of marginalized creators. Studies show that AI models trained to detect hate speech may perpetuate racial bias.^[44]

Following Red Dress Day on May 5, 2021, a day intended to raise awareness for Missing and Murdered Indigenous Women and Girls (MMIWG), Indigenous activists and allies found posts about MMIWG had disappeared from their Instagram accounts.^[45] Similar concerns were raised by Black Lives Matter advocates whose accounts were repeatedly flagged by Facebook, while Anti-Black racism was allowed to carry on without consequence. A participant commented that earlier in the year "when #Palestine was trending on Instagram, [the app] would shadow ban any users that wrote out the full word Palestine." This meant that any discussion or support for the topic was subdued.

Despite platform efforts to prevent and end online hate^[46], their underlying mechanisms prioritize profit over public interest. One participant called Instagram's safety features counterproductive, another added that hate speech filters are inherently oppressive. They observed that platforms only enforce rules against the most obvious and undeniable forms of hate (e.g., slurs), while accounts and posts targeting smaller, less-mainstream creators are left unchecked. This drives many marginalized and equity-deserving users off platforms or forces them to filter messages. Restricting message requests and other features causes women, gender diverse people and other equity-deserving creators to miss out on business, paid engagements, and networking opportunities.

Furthermore, moderation through natural language processing and human moderators tends to focus on major communities and languages, with little expertise or resources allocated to understanding the nuances of what is considered offensive in Indigenous or immigrant communities and other contexts. Given these issues, there is a real need to rethink the design, governance, and regulation of online social platforms rather than a narrow focus on limiting hateful expression in these spaces.



WHOSE FREEDOM?

The need for further protections to ensure safe online experiences for everyone, particularly youth and people from historically and systematically marginalized groups, is urgent work. However, the idea of hate speech legislation or hate speech provisions in human rights legislation is sometimes criticized for censorship[47], placing limitations on free and open democratic debate, and infringing on civil liberties. [48] It is, in fact, free and unfettered expressions of hate speech that undermines free speech, debate and democracy and causes physical and psychological harm for women, girls, gender diverse people, or members of Indigenous, 2SLGBTQIA+, Black, racialized, immigrant, or other equity-deserving communities. This 'Hate Speech versus Free Speech' debate must not be used as a rhetorical device to suppress any notions of government restrictions on hate speech.

Governments have a duty to prohibit hate speech while protecting diverse and dissenting opinions. Freedom of expression does not defend the freedom to speak in ways that attack, discriminate or incite violence and hatred.[49]

Currently, women, girls and gender diverse people do not enjoy freedom of speech because they are silenced, shut down and driven away from digital platforms by others whose exercise of free speech entails hate speech, intimidation, and threats. If governments fail to act, whose freedom of speech and expression are we protecting—the ones perpetuating harm or the ones being harmed?

The Canadian justice system and federal and provincial governments recognize that the right to express oneself, alongside other rights and freedoms enshrined in the Canadian Charter of Rights and Freedoms, are not absolute. They are subject to reasonable and justifiable limitations necessary to upholding the rights to equality, freedom from discrimination, and the core values underlying freedom of expression.[50] Proportional limits are required to prevent public dissemination of hate and protect diverse perspectives without systematically excluding any groups or voices. In essence, freedom of expression requires freedom from some of the most egregious and harmful expressions of hate. The government of Canada must take a balanced and discerning approach to safeguarding the right to expression while ensuring already marginalized groups are not paying a steep price for their participation on digital platforms.

"Currently, women, girls and gender diverse people do not enjoy freedom of speech because they are silenced, shut down and driven away from digital platforms by others whose exercise of free speech entails hate speech, intimidation, and threats."

GENDERED NATURE OF ONLINE HATE

Online social platforms are often hostile spaces, particularly for women, girls, gender diverse people and those whose intersectional identities leave them at heightened risk of experiencing online hate. Focus group participants disclosed that the hate they experienced or witnessed online was frequently and clearly gendered. One participant shared that “simply existing online as a woman, trans or gender diverse person is seen as an invitation to hate.” Others revealed that their gender, gender expression, gender identity or perceived gender combined with other factors such as sexual orientation, indigeneity, racialization, disability status, and age to make them a focus of hate online.

“Simply existing online as a woman, trans or gender diverse person is seen as an invitation to hate.”

- focus group participant

While people of all genders experience online hate, women and gender diverse people are disproportionately targeted in the digital arena. They experience substantially more hate online, both in terms of severity and sheer volume, simply for existing outside of the gender binary, for pushing back against rigid gender norms, or for being vocal about gender justice issues and matters of feminist and political concern.[51] Yet, online platforms systematically fail to recognize and respond to the severity of online hate directed to women and equity-deserving groups. A 2022 study by the Center for Countering Digital Hate (CCDH) analyzing thousands of direct messages of five high-profile women on Instagram found that the platform did not act on 90 percent of abuse sent via direct message to the women, despite the messages clearly violating community guidelines and being reported to moderators.[52]

The contours of online hate and harassment are undeniably and conspicuously gendered. Globally, over 7 in 10 people who report online hate are women or girls.[53] Among them, young and racialized people are most at risk. A poll commissioned by Canadian Race Relations Foundation (CRRF) through Abacus Data indicated that racialized Canadians are three times more likely to experience online hate speech as compared to the general population and that the age group most at risk is young people between 18 to 29.[54] Online hate is therefore a gender issue, and it is an intersectional issue.

Hate speech constitutes a form of gender-based violence (GBV) and is a way that misogyny and sexism operate online. Online hate is weaponized against women and gender diverse people to attack appearance, invalidate identity, endanger safety, police personal life, question values, shame sexual behaviour, damage reputation, silence resistance, and gatekeep online spaces. The digital expressions of gender-based violence, or technology-facilitated gender-based violence, include hate speech as well as stalking, sexual harassment, and non-consensual sharing of images.

There are clear gendered patterns in the hate speech directed at women and gender diverse people.[55] Focus group participants mentioned that even when their race, religion or other identity factors appeared to be the main reason for targeted hate, gender was an aggravating factor. According to a September 2022 national survey conducted by Environics Research and commissioned by YWCA Canada, the most common forms of online hate experienced by women and gender diverse youth 16-30 in Canada are sexist and misogynist in nature or based on body type or physical characteristics. [56] For women, girls, and gender diverse youth, revealing their gender identity meant receiving threats of sexual or physical violence from strangers slinging rape or death threats or making jokes and jabs of a sexual and discriminatory nature.

One participant spoke to the rise of gendered online hate, “now that pronouns are more prominent and their use is paid more attention to, this has given rise to a lot more hate speech around gender and sexual identity.” Several focus group participants noted a significant decline in experiencing hate online when they didn’t post their pronouns, pictures, or other markers of gender and sexual identity online. This points to another dimension of violence associated with online hate– digital repression and the erasure of self and identity. Elimination of hate speech is essential to ensuring that people of all genders and identities can fully participate in online platforms free from fear.

The most common forms of online hate experienced by women and gender diverse youth 16-30 in Canada are sexist and misogynist in nature or based on body type or physical characteristics.

DEACTIVATING DEMOCRACY: THE CHILLING EFFECTS OF ONLINE HATE CIRCULATION

A free and functioning democratic society relies on open and diverse public engagement. Unregulated and ineffectively moderated hate undermines democracy. It serves to exclude certain voices from online spaces and public life and poses a threat to community safety and social cohesion. Online hate against women and gender diverse people has a chilling effect[57] on democratic discourse[58], and their ambition and opportunity to be politically active [59] or participate in profile-raising opportunities and public-facing professions such as journalists, politicians, human rights defenders, activists, and artists. [60] Even when they do participate, their participation is precarious. Organized hate campaigns, gendered disinformation[61], and online threats and violence have been used to discredit and intimidate prominent public figures in Canada, particularly those that are women or from Black, Indigenous, racialized or 2SLGBTQIA+ communities. This includes Brandi Morin, Christine Labrie, Joanna Bernard, Rana Bokhari, Iqra Khalid, Rachel Gilmore, Erica Ifill, Chrystia Freeland, Catherine McKenna, Kathleen Wynne, Saba Eitizaz, and Raisa Patel among others.[62] The proliferation of online hate and lack of protections communicates that women and gender diverse individuals are undeserving of political office, public roles, and taking up space in the digital realm.

"Unregulated and ineffectively moderated hate undermines democracy"

Online hate is an obstacle to full and meaningful participation in society, leading to disenfranchisement and disengagement. The young women and gender diverse youth who participated in our focus groups maintained that experiences of online hate directly impacted their freedom of expression. Many responded to these occurrences by minimizing or limiting their online presence and participation in public life. As a result, young women and gender diverse youth withdraw from expressing themselves or exerting influence in digital spaces.[63] One participant mentioned the discrepancy in hate received by Indigenous-presenting content-creators in comparison to their white-passing counterparts. In this way, online hate contributes to the erasure and suppression of Indigenous cultures and heritage.

Our research shows that online hate or technology-facilitated gender-based violence does not have to be directly experienced to generate caution and self-censorship. The fear and apprehension from witnessing others encounter online hate dissuades new generations of women, girls, and gender diverse youth from engaging in public life and political office. The disenfranchisement of these groups impedes the diverse participation, representation, and leadership necessary for inclusive communities and civic engagement.

HOW YOUTH ARE RESPONDING

Beyond the staggering statistics and shocking headlines, the realities of online hate in relation to young people are under-discussed and unevenly experienced. The unprecedented digital shift of social life, work, and education—accelerated by the COVID-19 pandemic—has catapulted the use of internet services, content platforms and streaming sites. While the world is more connected than ever, this digital surge has put more young people in the line of harm and human rights abuses linked to the use of digital technology, with devastating impacts.

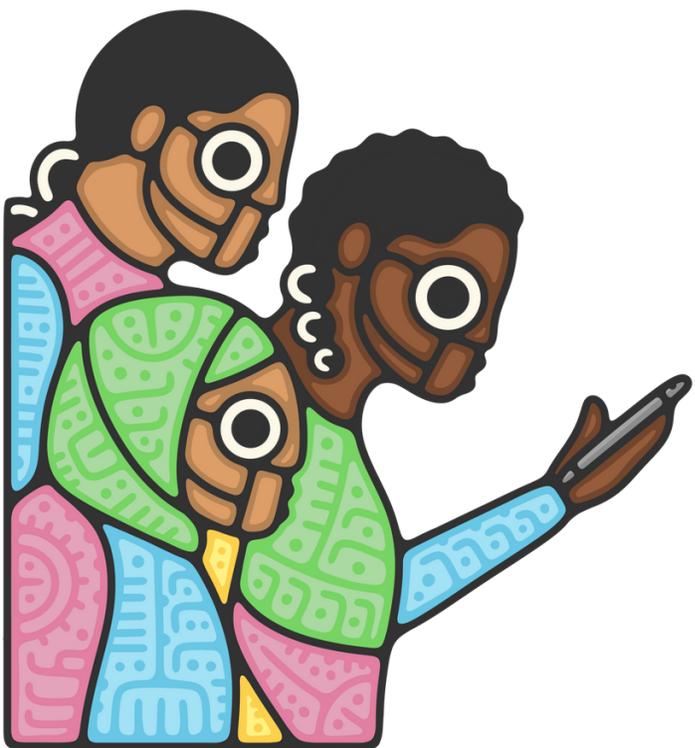
Without exception, every single focus group participant had recently experienced and/or witnessed some form of online hate or harassment.

Identities are in no small way socially constituted. What does it mean for self-formation and self-confidence when youth are increasingly thrust into online spaces where their dignity, values and existence are attacked? What strategies do they employ to respond to and speak up against hateful acts and attitudes?

01 Protecting
Themselves

02 Protecting
Others and
Organizing –
Active online
bystanders

Without exception, every single focus group participant had recently experienced and/or witnessed some form of online hate or harassment.



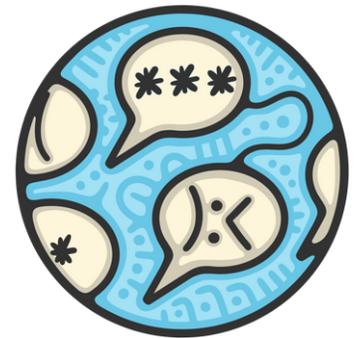
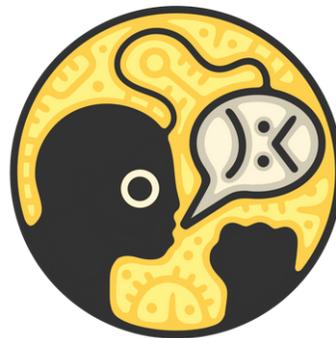
1. Protecting Themselves

The movement of online hate speech from fringe channels to mainstream online spaces has altered the nature of online engagement. Many participants reportedly resorted to self-preservation tactics such as filtering or restricting comments on their pages or posts to minimize exposure to online vitriol. Increasingly, the experiences young women, gender diverse youth, and marginalized groups are having online are leading them to self-censor, limit their digital interactions and, ultimately, leave online platforms.

Celebrities such as Elliot Page, Selena Gomez, Lesley Jones, Daisy Ridley, Ruby Rose, Chrissy Teigen and Lizzo have all talked about taking breaks from or quitting online platforms as a result of the damage done to their mental health and wellbeing. While taking a social media hiatus may seem like a good idea—and indeed it is—the consequences of disengagement are more complex. As online spaces become more toxic and hostile for women, gender diverse and marginalized people, they also become less accessible and representative. Individuals and groups that are unfairly and routinely targeted may disengage in online and offline public spaces or make their account settings private. Several participants mentioned not posting their face or their pronouns publicly. Others felt politically stifled.

The fear and stigma of online hate can impact how youth engage in any setting and keep them out of the public sphere, including running for office, becoming a journalist, volunteering in their community, or taking on leadership and visible roles. In these and other ways, hate speech can deny youth access to public speech. A participant mentioned that she was forced to cancel a human rights protest she was organizing because a radical group became aware of it and threatened to harass attendees.

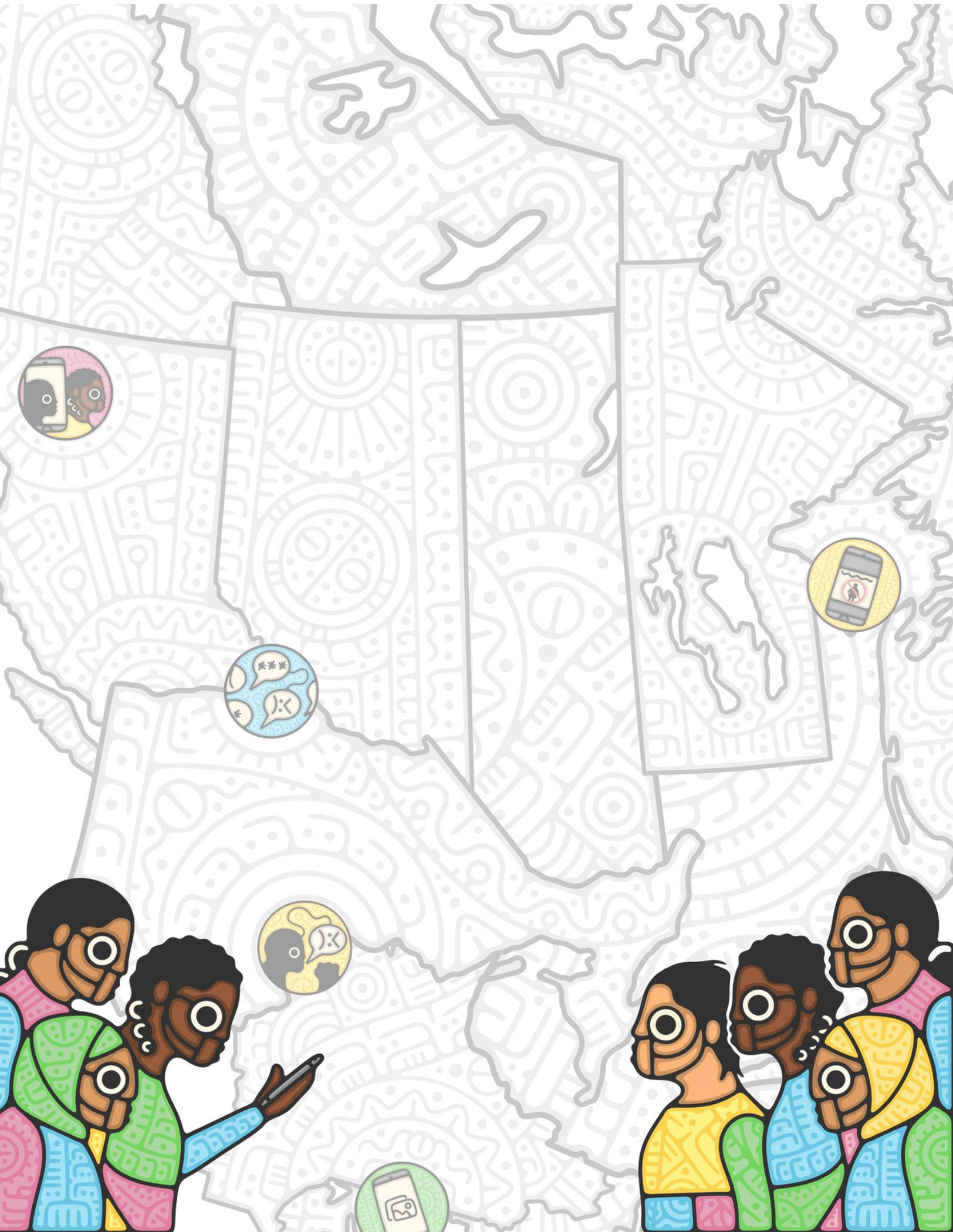
Participants who are gamers mentioned that online gaming forums and communities are 'hotbeds of hate'; with little attention to moderation and protection.[64] Participants who are gamers mentioned that online gaming forums and communities are 'hotbeds of hate'; with little attention to moderation and protection.[65] and "swatting"[66] can lead to death due to ongoing and historical police violence. The question remains, what parts of themselves are they are having to repress, to hide, to minimize to feel safe online...?



2. Protecting Others and Organizing – Active online bystanders

The youth engaged in our focus groups were largely attuned to the implications of online hate and the need for coordinated action. Many were vocal allies of marginalized or maligned communities and used their platforms to speak up in defiance of hate-filled rhetoric. Respondents highlighted the strategic possibilities of internet platforms for organizing and public education. Several mentions were made of queer, fat, trans, disabled, Black, Indigenous, and/or women of colour pages, accounts and public figures that center affirming, anti-oppressive ideas. Participants shared examples of how young celebrities and activists such as non-binary activist Alok Vaid-Menon modelled compassionate, authentic and alternate ways to respond to online hate. Just as the internet is used as a vehicle of hate, some traced the genealogy of powerful social movements like Black Lives Matter and the Me Too movement to digital responses to discriminatory behaviour. However, it was widely acknowledged how easy it was to get caught up in online herd mentality and engage in hateful behaviour or bullying. When hate is normalized, the fear of social exclusion or alienation holds many internet users back from speaking up in defense of themselves or others, and taking action against hate in all its forms.





04 Calls for Action from Youth Consultations Across Canada

Hate that manifests online cannot be delinked from hate that occurs offline. Both have real consequences and are deeply connected. The implications of online hate on psycho-social wellbeing, human rights and public safety are immense, warranting intersectional interventions. Research has shown that trolling, sexually explicit emails or messages, cyberbullying, rape threats, non-consensual sharing of intimate images, cyberstalking and doxxing are part of a continuum between online and offline manifestations of technology-facilitated violence against women, gender diverse and equity-deserving communities.[67] Participants mentioned the lasting trauma and recurring fears associated with online hate. They want to ensure online hate is taken seriously in and of itself and not only because it can be linked to or lead to offline hate.

The content and recommendations presented in this report center the voices, lived experiences, and potential solutions shared by young women and gender diverse youth aged 16-30 years.

YWCA Canada convened ten focus groups in February and March 2022, facilitated by our Research Team and Leadership Fellows, to discuss individual and collective experiences of online hate across Canada and develop community-generated solutions to curb the circulation of digital hate and mitigate its harms.

These recommendations include actionable steps for (i) Online social platforms, (ii) for the Canadian federal government, (iii) for Non-Profit and Community Organizations and, (iv) Youth and all users of online platforms. Crucially, participants called for a whole systems approach, involving coordinated collective action by multiple sectors and stakeholders to effectively respond to digital hate.

Recommendations and policy perspectives emerging from the consultations can broadly be categorized under four themes: Advocacy, Accountability, Support and Regulation.

Advocacy & Awareness

Accountability

Support

Regulation

ADVOCACY AND AWARENESS

“Power concedes nothing without a demand.” - Frederick Douglass, 1857

Among the key interventions articulated by our participants was the need for digital users and other stakeholders to confront the status quo and push for positive change. Public education and counterspeech are key starting points to uproot hate speech from digital spaces. Youth participating in our focus groups were full of ideas on how to confront online hate and effect change in online environments.

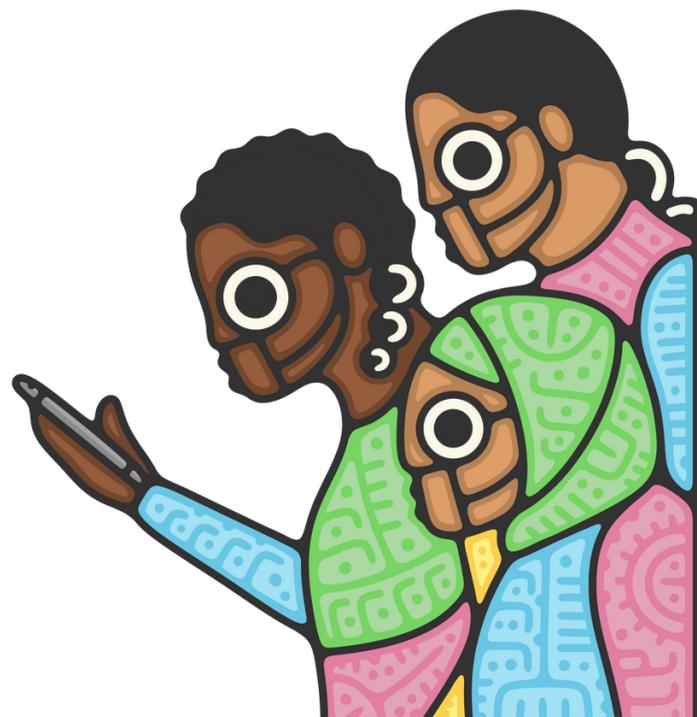
1. Leverage Critical Mass

“It is easy to say that social media companies are responsible, but we are the ones keeping them open. At the end of the day, these are community spaces...” Focus Group Participant

Part of the reason that current conditions are left unchanged or unchallenged, is that it is not widely known how many internet users are impacted by hate online and in what ways. Statistics and intersectional analysis can be an awakening. However, clear, reliable, and disaggregated data around the incidence, nature and responses to online hate are currently not make publicly available by online platforms. Participants advocated for statistics and resources that map the extent and nuances of the problem of online hate that could then form the basis of pro-active collective action through stakeholder mobilization.

Alongside bringing visibility to its prevalence, participants highlighted that banding together to push back can force online platforms to take notice of the demand for spaces free of hate speech. While participants pointed to examples of allies and advocates taking action against online hate, they acknowledged that a

coordinated effort that recognizes the strength in numbers can more effectively drive change. Social media companies are always gathering user feedback and monitoring user behaviour. Quitting, disengaging, or limiting online presence may not be the answer. However, participants raised that if enough users speak up or migrate en masse to apps or platforms that center safety and online wellbeing, companies will take notice and make changes to their terms of service and policies around hateful conduct and harassment. Coordinated campaigns can foster public solidarity and support for victims of online hate and spur further action.



2. Digital Literacy, Safety and Resilience – Online Bystander Training

“People join socials every single day and a lot of people don’t know how dangerous it can be. It’s important to help share information and education on best ways to use socials for the reason you got it for, [how] to protect yourself, how to follow the best people, how to get the algorithm to show you content you actually want to see, [and] being intentional about social media usage...there are options to block, mute, unfollow – use them!” Focus Group Participant

Digital, media and information literacy play an important role in developing a person’s technical and critical skills.[68] A systematic response to online hate must equip and educate people to identify, respond to, and protect themselves from hate occurrences online. Participants highlighted the need for mandatory sensitivity and safety training for both platform users and platform employees. Partnerships between civil society, community members, governments, and private technology companies are key to implementing an intersectional gender equity lens to analyze and address online hate.[69]

Given how early children are getting online, how often they encounter digital hate, and how easily hate starts to take root, education interventions must be offered early and often.[70] Issues of identity and isolation that young people grapple with can render them more vulnerable to hateful messaging or groups.[71] Childhood and adolescence are crucial points of intervention in preventing online hate speech.

teaching kids about empathy, anti-racism and anti-oppression, critical thinking, and positive relationship-building can encourage respectful and safe online conduct and empower them to resist and report hate speech.[72] Classrooms can prepare children to deconstruct cultures of hatred[73] and become agents of change. The anti-hate toolkit for educators by Canadian Anti-Hate Network, “Confronting Hate in Canadian Schools,” provides real-world, practical steps for students, educators, administrators, parents, and community members.

Non-Profit and Community organizations are a key node in building and disseminating fundamental tools, training and resources. In collaboration with the Anti-Hate Community Leaders’ Group, the Council of Agencies Serving South Asians (CASSA) developed a toolkit for organizations and victims in an effort to combat online hate through safety training, effective online allyship and bystander intervention [74]. Public campaigns to cultivate tolerance and confront hate can help crack down on hate speech and spread awareness of affirming and alternative counter messaging that dilutes online hate. The European No Hate Speech movement has organized training sessions for bloggers and young activists to share best practices in implementing practices that uplift and portray marginalized communities in a positive way and deflect hateful discourse. [75]



3. Offline efforts to uproot oppression and drive systemic change

Creating a coordinated strategy to eliminate hate speech begins with the recognition that online hate is a reflection of offline systems of oppression and marginalization.[76] To tackle the rise and spread of online hate, we must address its offline origins. Racism, misogyny, patriarchy, ableism, homophobia, transphobia and other corrosive forces of discrimination are deeply embedded structural elements of society. When these hateful agendas find expression online, they expose more people to harm and exacerbate the insidious impacts online and offline. Participants gestured to multi-stakeholder approaches to dismantling offline systems of oppression as key to disrupting online manifestations of hate.

4. “Heart” Speech: Alternative counternarratives

The Internet is complex. It can be educational, empowering, fraught, and often dangerous for young people, especially young women, gender diverse youth, racialized people and other equity-deserving groups. While content takedown and account blocking approaches to online hate are necessary, they cannot be the single solution, given the scale of the problem. Further, these approaches do little to mitigate lasting harms, educate perpetrators and viewers, or empower survivors to speak back or speak up about these issues.[77] Survivors who spoke up in our focus groups mentioned the hurt that lingered long after hateful posts or content were removed.

As part of the larger toolkit against online hate speech, online counterspeech can be a powerful and restorative way to reclaim online spaces by offering alternative viewpoints, centering survivor voices, and creating opportunities for community education and collective healing.[78]

Online counternarratives can be a powerful form of digital feminist response and resistance. With platform delays in resolving reported content and taking down hateful and offensive posts, users are increasingly leaning on their own labour and wits to talk back to ‘haters’ and draw attention to the patterns of hateful messaging and abuse online. One participant shared the inventive way they were deflecting transphobia online. When confronted with a statistic that is often used by hate instigators in comment sections to dismiss trans identities, diminish their existence and encourage self-harm by trans people, they responded with a sharp rebuttal “oh you mean the number of police officers that are implicated in domestic violence cases?”.

Youth participants strongly believed that their responses to perpetrators and discourses of hate can help to de-escalate and diffuse situations and develop new ways of online engagement. They directed us to public figures using their platforms to ‘clapback with kindness’ and draw attention to digital harms. How can alternative narratives help them negotiate and reclaim online spaces? What kinds of positive digital counter publics can young people find, or create, in an environment that is fundamentally violent and exclusionary?

ACCOUNTABILITY

An overarching theme emerging across the focus groups was the acknowledgement that online social platforms are not passive distributors of user-generated content. Platforms actively promote and profit from the circulation of hate through algorithms that amplify divisive content and deprioritize marginalized or racialized perspectives. In fact, evidence has shown that online spaces and their content recommendation systems can effectively function as pipelines for polarization and rabbit holes of radicalization.[79] For instance, when social media companies do not act to weed out hate and extremist ideologies such as the “Great Replacement Theory”[80], it makes white supremacist, neo-Nazi theories more accessible and acceptable to a wider, susceptible audience. Events as recent as the May 2022 mass-shooting in Buffalo, New York show how exposure to racist rhetoric and activity online can in turn exacerbate racist violence and hate crimes. Platforms taking content moderation obligations seriously can help ensure safer online environments that supports participation. In addition to improved moderation, participants called for platforms to take down accounts, communities, forums, subreddits [81] and bots[82] ‘dedicated to the dissemination of hate’.

Online social platforms should take accountability for the messages and values that find a home on their platforms. Participants argued that the revenue generation model for social media companies relies on increased engagement and controversy. This inherently makes companies change-resistant since taking away hate means taking away income. While they may be part of the problem, online platforms can be a part of the solution and a force for progressive change. Legislative efforts and advocacy campaigns can urge platforms to adjust their approach to online hate speech.

1. Proactive Preventive Measures

Current content regulation and online reporting mechanisms do little to reduce the proliferation of online hate. A number of participants mentioned that the use of bots or machine-learning instead of people to moderate allows for gaps in identifying and responding to perpetrators of online hate. On the flip side, reliance on human moderators causes excessive backlogs and exposes more people to hate.

Delays in responding to reported accounts or content further rendered these avenues ineffective or less likely to be utilized. Critical to protecting youth from hate-motivated behaviour online, but largely absent from current strategies, are preventive strategies to address the systemic root causes of hate speech. Rather than reactive measures that do little to reduce hateful acts, several participants called for approaches that seek to prevent and deter users from engaging in hate-based speech or behaviour.

Mandatory anti-oppression and sensitivity training for employees[83] and users, as well as requiring a signed code of ethics agreements when signing up to social media was suggested as a way to help hold people accountable to not committing hate speech. Civil society organizations can reinforce this message through educational outreach and governments can follow through with clear and consistent online hate regulations.

2. Innovative Technological Interventions

Leveraging new and existing technology, in consultation with community members and organizations, may yield innovative tools to counter online hate. Participants pointed to content moderation features introduced by companies to curb misinformation and/or encourage respectful communication. Meta has piloted 'warning screens'[84] that appear after someone has typed a post, to educate users and discourage hate speech, harassment and content that may violated community standards from being posted. Twitter's suite of measures include pop-up notifications when a user attempts to retweet a "disputed tweet", alerts when participating in discussion around potentially "volatile" or "intense" topics, prompts to review messages that could be "harmful or offensive", and reminders to engage respectfully. Tumblr, Pinterest and Instagram have policies and Public Service Announcements around content flagged for promoting self-harm or extremism.

New measures inviting users to reflect or redirect before sharing or engaging with hateful or harmful content were welcomed. Suggestions from participants included interventions that share information on support resources, alternative content recommendations, and "kind comeback generators"[85] to intercept online hate. The need to consult with community members experiencing hate was highlighted as a fundamental step towards developing technology tailored to create safer, more positive online spaces.

3. Transparent and Collaborative Processes

Our discussions yielded the idea that there was very little report-back, once digital users hit the 'report' button. People flagging content or an account as hateful or against community guidelines, were often not contacted or given an explanation once review was conducted. There is very little disaggregated information publicly available around the nature and scope of hate incidents occurring on different online platforms, and how this is being dealt with. Community groups can mobilize youth to campaign for more openness and transparency.

Part of ensuring that online spaces are safe and accessible is ensuring that they are representative and responsive. While online platforms have established Departments of Community Development and Directors of Trust and Safety, little is made known about their composition and process.

"The need to consult with community members experiencing hate was highlighted as a fundamental step towards developing technology tailored to create safer, more positive online spaces."

One participant called for platforms to have more people in executive positions who “solely focus on keeping these platforms safe”. Another added that it was important to see people with intersectional identities from marginalized communities involved in content moderation and policymaking on hate regulation. Bringing radical transparency into company reporting can restore trust in the efficacy of processes and allow users to make informed decisions about their continued engagement with a platform and its community. [86] Efforts must be made to consult and involve young people as closely as possible in (re)framing community guidelines and content policies.



SUPPORT

01 Better Resource Allocation for Safer Spaces and Supportive Interventions

02 Community-Centered Online Platforms

1. Better resource allocation for safer spaces and supportive interventions

The capacity to create change is often contingent on funding cycles and approval. As active members of YWCA member associations and other community organizations, our participants were cognizant of the fact that grassroots efforts to tackle online hate were undermined by insufficient funds or complex grant processes. Long-term investment and support for digital innovation is required to creatively and effectively confront online hate.

Participants called for streamlined processes to fund community organizations actively responding to online hate and developing supports for those experiencing online harms.

Furthermore, this push would encourage companies to allocate more funds towards technological interventions for safer cyberspace, community engagement, as well as training and therapy supports for content moderators who sift through the most violent and vitriolic content on the internet.

2. Community-centered online platforms

Participants enthusiastically pointed to a small crop of community-focused apps and websites that place safety and positivity at the core of their business model. While these options are welcome and necessary alternatives, greater support is required to bring these changes to mainstream channels. Some young women and gender diverse youth in our focus group cohorts reported turning more to networks that foster safer, kinder community, this gives more space for hate to fester in mainstream platforms. To make an impact at scale, community-based media forums must receive adequate funding and resources. [87]

REGULATORY RESPONSES TO ONLINE HATE

01 Canadian Legislative Context

02 Calls for Feminist Federal Action

1. Canadian Legislative Context

Although online hate is generally a slippery target for legal instruments, the Canadian policy context makes addressing online hate more achievable through potential legislation and regulatory frameworks. In Canada, anti-hate laws at the federal, provincial and territorial levels impose reasonable limits on the freedom of expression guaranteed by section 2(b) of the Canadian Charter of Rights and Freedoms. [88] Sections 318 and 319 of the Criminal Code impose criminal sanctions against anyone who advocates genocide, publicly incites violence and willfully promote hatred against an “identifiable group.”[89] A landmark Supreme Court of Canada judgement in 2013 ruled that laws against hate speech are a reasonable limit on freedom of expression and justifiable by the Charter. [90] Other human rights laws aim to prohibit the publication of discriminatory messaging that perpetuates dangerous propaganda or targets members of particular groups.

A recent poll commissioned by the Canadian Race Relations Foundation found that three quarters of Canadians would like to see more government action to combat hate and racism online and nearly four in five Canadians (79 percent) support the introduction of legislation to combat serious forms of harmful online content, including hate speech. [91]

79 percent of Canadians support the introduction of legislation to combat serious forms of harmful online content, including hate speech.

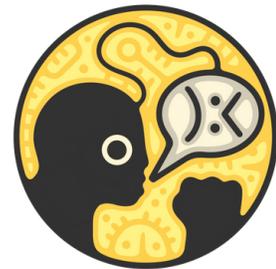
Acknowledging the need for clear and targeted action, the Government of Canada announced a coordinated approach to developing a legislative and regulatory framework to address online harms. The 2021 Liberal party platform promised urgent efforts to “combat serious forms of harmful online content, specifically hate speech, terrorist content, content that incites violence, child sexual abuse material and the non-consensual distribution of intimate images.” [92]

Bill C-36 proposed to amend Canada's Criminal Code to codify the definition of hate speech, revive section 13 of the Canadian Human Rights Act that makes it a “discriminatory practice” to communicate hate speech over the internet, and improve the complaint and redressal process for victims. However, it has been noted that the proposed framework could exacerbate the “existing, well-documented pattern of online speech policing and removal targeting equity-deserving individuals and communities”. The speed at which proposed legislation required platforms to take down or resolve reported or flagged accounts and content, would necessitate the use of machine learning and Artificial Intelligence. This reliance could perpetuate algorithmic bias and inequalities.

Community organizations and experts are actively engaging in consultation with and providing recommendations. This is a crucial moment to build on this momentum and mobilize youth for change.

2. Calls for Feminist Federal Action

Real progress on eradicating online hate requires systems level efforts led by government in collaboration with civil society and online platforms. Participants called for intersectional government intervention to engender greater accountability and ensure online platforms are legally accountable for their users’ online safety and liable for harms that occur on or are facilitated by their platforms.





05 Feminist Regulatory Responses to Online Hate - A Federal Responsibility

This report provides fundamental considerations and specific recommendations for safer online spaces emerging from research and community consultations. It offers five key principles that all levels of government must consider as they develop and implement policy and legislative responses to online hate. Building from that, it proposes fifteen recommendations for federal legislative action and commitment organized around five pillars.

Five fundamental considerations:

1

The Government's Role in Online Safety

The federal government plays a key role in promoting online safety and addressing hate speech in digital spaces, and should introduce regulations and implement legal reform to confront online harms. This includes enacting a legal obligation and liability framework for online platforms and establishing a regulatory entity to monitor and evaluate implementation.

While there is a role for federal, provincial, territorial, and municipal governments in online safety, recommendations in this report apply primarily to the Canadian federal government and may require intergovernmental collaboration.

2

Intersectional and Survivor-Centric Approach

The perspectives, needs and solutions of survivors and directly impacted communities should be at the forefront of this framework, with the explicit goal of creating safer online spaces for all Canadians and particularly women, girls, gender diverse people and those whose intersecting identities make them disproportionately vulnerable to online harms. Legal reforms and platform regulations must center community-generated solutions, trauma-informed approaches, and intersectional feminist considerations.

3

Recognition of Individual and Systemic Harm

The regulatory framework must recognize that in addition to causing individual harm, online hate also constitutes systemic harm to communities, particularly towards historically marginalized groups and equity-deserving groups – impacting their ability to participate fully in society and meaningfully exercise their human rights.

4

Guarantee Right to Privacy and Participation

The framework must empower platform users and ensure the right to privacy, freedom of expression, and safeguard the ability of all people to participate fully and meaningfully in digital discourse and public life. The Canadian legislative context acknowledges that implementing proportional limits to freedom of expression may be necessary to uphold the rights to equality, freedom from discrimination, and freedom of expression of individuals and communities that are disproportionately impacted by online harms.

5

Definitional Clarity and Legislative Scope

The regulatory framework must have a clear focus on online hate and related harms. Given the wide range of forms and expressions of online hate, statutory language must clearly define online hate in the context of online platforms and set out the scope and intention of legislation. While disinformation and digital privacy are other critical areas of attention, this legislation must make targeted efforts to curb online hate and not risk diluting efforts with a wide, ill-defined scope.

Recommendations for federal action have been categorized under five broad pillars: (i) Legal Framework, (ii) Platform Regulation and Content Moderation, (iii) Oversight, Transparency and Accountability, (iv) Intersectional Approaches to Addressing Online Hate, (v) Research, Education and Support.

LEGAL FRAMEWORK

1. The law should define and delineate which online platforms and services are covered by regulation and create categories based on the nature of content and potential harms. Recommendations included in this report speak directly to regulation of online social platforms and do not include internet infrastructure providers (e.g., Bell, Telus, Rogers), electronic mail (e.g., Gmail, Outlook, Yahoo), voice and video communication services (e.g., Skype, Zoom, Google Meet/Duo, FaceTime) and private messaging applications (e.g., WhatsApp, WeChat, Telegram, Signal) within their scope.

2. Online platforms and services such as social media and video sharing sites should be viewed as unique in the ICT landscape and be subject to regulatory requirements that recognize their unique features and the distinct role they play in Canadian society and in facilitating the spread of online hate speech. In addition to broad, baseline regulations, legislation must include proportional and specific obligations for different categories of online platforms that recognize the variability in user base, content, and risk of harm.

3. Legal approaches to online harm mitigation must include the regulation of hate speech in addition to hate crimes. The regulatory framework should address not only clearly illegal content, but be flexible enough to respond to harmful content, recognizing the detrimental impact that this content can have on individuals and communities. Further, the legal definition of online hate in Canada must be reviewed to ensure it is responsive to the evolving nature of hate speech and behaviours online.

Development of coherent terminology has practical implications for legally obliging platforms to prohibit hate speech and not leaving it open to interpretation by platforms and content moderators.

PLATFORM REGULATION AND CONTENT MODERATION

4. Ensure regulatory coverage of public-facing features and private messaging functions of online social platforms. Private messaging features on these platforms (such as Instagram DMs, Twitter DMs and TikTok DMs) extend the possibility of hate occurrences, and as such should not be excluded from legislative purview.

5. The regulatory framework must establish a set of requirements from online social platforms, including at minimum:

- Proactive and preventative measures including community safety standards and features or prompts to give users pause before posting potentially harmful content
- Mandatory anti-oppression, sensitivity, and safety training for staff
- Accessible and expedient user reporting and redressal mechanisms
- Clear and effective complaint review and content moderation processes
- Country-specific and function-specific reporting of incidents and their resolution



6. Designation of an ombudsperson or department at the platform level with governance and oversight functions to ensure content moderation, removal and/or reinstatement, and data sharing are conducted from a survivor-centric, trauma-informed and harm mitigation approach. Building, monitoring, and updating content moderation systems that combine Artificial Intelligence (AI) automation or machine learning with human expertise could be an effective industry standard to adequately and appropriately address bias in algorithms, regulate online hate, and create safe online environments.

OVERSIGHT, TRANSPARENCY AND ACCOUNTABILITY

7. The government should create a centralized quasi-governmental agency with a public mandate to protect and promote online safety. This Digital Safety entity would be tasked with regulatory oversight of the online safety regime. Strict protocols should be put in place to ensure that the Digital Safety regulator operates with accountability while acting independently of government. Relative autonomy of operations is essential to ensure the entity remains non-partisan and holds the line to ensure placement of online safety obligations on online social platforms do not lend to state control, surveillance, or repression of online expression. Regular monitoring of regulatory decisions and arbitration on sensitive issues and matters of national security concern is required to maintain compliance with Canadian law and human rights frameworks.

8. The Digital Safety regulator should be provided with authority to evaluate platform policies and enforce mechanisms to guarantee online social platforms comply with their obligations

under the legislation, including the ability to apply sanctions and levy fines for non-compliance. This body must also be empowered to handle user appeals and arbitrate resolutions for contested content moderation decisions at the platform level. The Digital Safety regulator's resolution and remedial processes should be made available to individuals that have already gone through the platform's internal process as well as those that have not used them.

9. Online social platforms should be required to act responsibly, with transparency, and be held accountable for the risks their services and processes can pose to society, democracy, and citizens. Online platforms must be obliged to:

- Collect, report, and make publicly available, data on online hate that is reported and addressed on their platforms. They must ensure that data collected is disaggregated and intersectional, in order to understand how women and gender-diverse people, racialized people, 2SLGBTQIA+ communities and other marginalized groups are impacted by online hate.
- Maintain transparency in how content that is reported or flagged is reviewed. Share the framework for making decisions on this content and publish comprehensive annual transparency reports to show how platforms are addressing online hate.
- Platforms must make timely and transparent reports back to individuals who reported or flagged content with the outcome and the reasons for the decision.
- Undergo independent audits (conducted by the Digital Safety regulator).

INTERSECTIONAL APPROACHES TO ADDRESSING ONLINE HATE

10. Any regulatory framework or government action on online hate must recognize online hate as real harm and acknowledge the connections between online hate and offline violence.

Individuals impacted by online hate experience real-world harms Any regulatory framework or government action on online hate must recognize online hate as real harm and acknowledge the connections between online hate and offline violence. Individuals impacted by online hate experience real-world harms and can experience profound and lasting psychological effects. Policy makers and platforms must take online hate seriously in and of itself and not only because it can be linked to or lead to offline hate.

11. Legislation alone cannot solve the issue of online hate and related harms. Meaningful government action against online hate requires multistakeholder approaches to address the root causes of online hate, namely white supremacy, homophobia, transphobia, misogyny, colonialism, ableism and other intersecting systemic oppressions in Canadian society. For instance, civil society agencies must work with government to triage and disperse resources to impacted communities in the wake of intensifying online hate occurrences.

12. Legal instruments must recognize online hate is often an expression of gender-based violence and apply a Gender-Based Analysis (GBA+) lens to understand the intersectional impacts of legislation and policies and analyze how equity-deserving communities can be protected from consequences of new regulatory frameworks:

- Online hate legislation must be harmonized with the scope and substance of a National Action Plan to end gender-based violence.
- The offices of the Digital Safety regulator and platform-level ombudsperson must be staffed and managed by individuals with expertise in understanding and tackling social injustices and technology-facilitated violence from an intersectional lens.
- Platforms' internal reporting or remedial processes and those of the Digital Safety regulator must provide appropriate options for individuals who do not want to be involved with or have been systematically oppressed by law enforcement or the criminal justice system.
- Platforms and the regulator must not be required to pass on information to the Royal Canadian Mounted Police (RCMP), Canadian Security Intelligence Service (CSIS), local or provincial police without the informed consent of the complainant, unless the content is evidence of a criminal offense or national security concern.
- Government must respect and understand that many internet users remain fearful or mistrustful of the legal and criminal justice systems and refrain from using reporting mechanisms or legal recourse. In focus groups, participants mentioned how practices like 'swatting'¹ can target Black, Indigenous, racialized, and other marginalized content creators and even lead to death due to ongoing and historical police violence.

RESEARCH, EDUCATION AND SUPPORT

13. Sustained funding should be provided to non-for-profit and civil society organizations working to build resilience against online hate and support survivors in communities across Canada. Resources must be dedicated to advance research, training, capacity-building, collaboration, community-based programming, public education, and peer-supports.

14. The mandate of the Digital Safety entity tasked with implementation, oversight, and enforcement of Canada's online safety regime should include a prevention, public education, and democratic engagement component.

15. Fund and make available community interventions including education resources, support programs, training and leadership, and 'carespaces' to discuss and collectively heal from hate incidents. Resource community organizations to offer context-specific supports, and youth-specific interventions including counselling services, digital literacy, security, leadership and advocacy programs opportunities to educate and engage youth in anti-hate leadership.





KEY REFERENCES

Khoo, C. (2021). Deplatforming Misogyny. Report on Platform Liability for Technology-Facilitated Gender-Based Violence. Women's Legal Education and Action Fund (LEAF). Toronto, Ontario. Available [here](#).

L1ght (2020). Rising Levels of Hate Speech & Online Toxicity During This Time of Crisis. Available [here](#).

Plan International. (2020). Free to be Online? Girls' and Young women's Experiences of Online Harassment. The State of the World's Girls Report. Available [here](#).

Posetti, J. et al. (2021). The Chilling: Global trends in online violence against women journalists. Research Discussion Paper. UNESCO. Available [here](#).

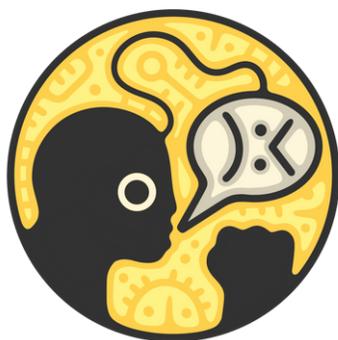
Public Policy Forum. (2019). Policy Approaches to Online Hate. Summary Report. Available [here](#).

Stevens, Y. and Krishnamurthy, V. (2021). Overhauling the Online Harms Proposal in Canada: A Human Rights Approach. Ottawa Faculty of Law Working Paper No. 2021-25.

The Center for Countering Digital Hate (CCDH) and Human Rights Campaign (2022). Digital Hate Report: Social Media's Role in Amplifying Dangerous Lies About LGBTQ+ People. Available [here](#).

The Center for Countering Digital Hate (CCDH). (2022). Hidden Hate: How Instagram fails to act on 9 in 10 reports of misogyny in DMs. Available [here](#).

UN Women– Asia Pacific. (2021). Eliminating Online Hate Speech to Secure Women's Political Participation. Action Brief. Available [here](#).



ENDNOTES

[1] The terms 'online social platforms', 'online platforms' and 'platforms' are used interchangeably in the report to refer to the distinct role that social media websites and internet-based applications (such as Instagram, Facebook, Twitter, Snapchat, TikTok and Reddit) play in advancing social connectedness and the spread of user-generated content and ideas. While online hate is present across a range of internet platforms, the scope of this report does not cover private messaging platforms (such as WhatsApp, Telegram or Signal), live-streaming platforms, e-commerce portals, news or blogging websites.

[2] Communities who experience barriers to equal access, opportunities, resources, and participation due to historical marginalization, structural disadvantages and/or social discrimination. These groups include include, but are not limited to, women, Indigenous peoples (First Nations, Inuit, and Métis), persons with disabilities, members of racialized groups and members of 2SLGBTQIA+ communities.

[3] Backe, E. L., Lilleston, P., & McCleary-Sills, J. (2018). Networked individuals, gendered violence: A literature review of cyberviolence. *Violence and Gender*, 5(3), 135-146.

[4] Hern, A. (2020). TikTok expands hate speech ban. *The Guardian*. Available at: <https://www.theguardian.com/technology/2020/oct/21/tiktok-expands-hate-speech-ban>

[5] Pandey, N., & Pal, A. (2020). Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. *International Journal of Information Management*, 55.

[6] Orlando, J. (2020). Young people are exposed to more hate online during COVID. And it risks their health. *The Conversation*. Available at: <https://theconversation.com/young-people-are-exposed-to-more-hate-online-during-covid-and-it-risks-their-health-148107>

[7] Environics Research and YWCA Canada (2022). *Experience with Online Hate: A national survey to better understand young women and gender diverse people's experiences with online hate in Canada. Report of Findings, September 8, 2022.*

[8] Brisson-Boivin, K. (2019). *Pushing Back Against Hate Online*. MediaSmarts. Ottawa.

[9] Orlando, J. (2020). Young people are exposed to more hate online during COVID. And it risks their health. *The Conversation*. Available at: <https://theconversation.com/young-people-are-exposed-to-more-hate-online-during-covid-and-it-risks-their-health-148107>

[10] Brisson-Boivin, K. (2019). *Pushing Back Against Hate Online*. MediaSmarts. Ottawa.

[11] L1ght (2020). *Rising Levels of Hate Speech & Online Toxicity During This Time of Crisis*. Available at: https://l1ght.com/Toxicity_during_coronavirus_Report-L1ght.pdf

[12] Lerman, K., Yan, X., & Wu, X. (2016). The "Majority Illusion" in Social Networks. *Plos One*, 11(2).

[13] CBC News (2017). *Canadians appear to be more hateful online. Here's what you can do about it*. CBC Marketplace. Available at: <https://www.cbc.ca/news/canada/marketplace-racism-online-tips-1.3943351>

[14] Environics Research and YWCA Canada (2022). *Experience with Online Hate: A national survey to better understand young women and gender diverse people's experiences with online hate in Canada. Report of Findings, September 8, 2022.*

[15] 2SLGBTQIA+ is an acronym that stands for Two-Spirit, Lesbian, Gay, Bisexual, Transgender, Queer or Questioning, Intersex, Asexual, and additional sexual orientations, and gender identities. The use of the acronym is intended to affirm the many ways people choose to self-identify their gender identity and sexual orientation and an invitation to encourage expansive understandings of sexuality, queerness, identity, rights, privilege, and colonialism.

[16] Canadian Race Relations Foundation (CRRF-FCRR) and Abacus Data (2021). Poll demonstrates support for strong social media regulations to prevent online hate and racism. January 25, 2021. Available at: <https://www.crrf-fcrr.ca/en/news-a-events/media-releases/item/27349-poll-demonstrates-support-for-strong-social-media-regulations-to-prevent-online-hate-and-racism>

[17] For a discussion of how misogyny and racism are often intertwined in online hate occurrences, see: Saresma, T., Karkulehto, S., & Varis, P. (2021). Gendered violence online: Hate speech as an intersection of misogyny and racism. In *Violence, gender and affect* (pp. 221-243). Palgrave Macmillan, Cham.

[18] Matamoros-Fernández, A., & Farkas, J. (2021). Racism, hate speech, and social media: A systematic review and critique. *Television & New Media*, 22(2), 205-224.

[19] Bliuc, A., Faulkner, N., Jakubowicz, A., & McGarty, C. (2018). Online networks of racial hate: A systemic review of 10 years of research on cyber-racism. *Computers in Human Behaviour*, 87, 75-86.

[20] TFGBV refers to a spectrum of activities and behaviours that involve technology as a central aspect of perpetuating violence, abuse, or harassment against (both cis and trans) women and girls. The phenomenon includes those who hold intersecting marginalized identities, such as 2SLGBTQIA+, Black, Indigenous, and racialized women; women with disabilities; and women who are socioeconomically disadvantaged. For a detailed consideration of online misogyny and TFGBV, see Khoo, C. (2021). *Deplatforming Misogyny. Report on Platform Liability for Technology-Facilitated Gender-Based Violence*. Women's Legal Education and Action Fund (LEAF). Toronto, Ontario. Available at: <https://www.leaf.ca/publication/deplatforming-misogyny/>

[21] For extensive discussion of gendered dimensions of online hate, see Adams, C. (2018). "They Go for Gender First" The nature and effect of sexist abuse of female technology journalists. *Journalism Practice*, 12(7), 850-869; Flynn, A., Powell, A., & Sugiura, L. (2021). *The Palgrave Handbook on Gendered Violence and Technology*. Palgrave Macmillan; KhosraviNik, M., & Esposito, E. (2018). Online hate, digital discourse and critique: Exploring digitally-mediated discursive practices of gender-based hostility. *Lodz Papers in Pragmatics*, 14(1), 45-68.

[22] The Online Hate Research and Education Project, a Government of Canada Anti-Racism Action Program funded initiative, has launched its anti-hate website. Resources made available on the website will help identify and contextualize the ever-changing nature of online hate and related harms. Available at: <https://hatepedia.ca/>

[23] See Davey, J., Hart, M., Guerin, C., & Birdwell, J. (2020). An online environmental scan of right-wing extremism in Canada. Institute for Strategic Dialogue. Available at: <https://www.isdglobal.org/wp-content/uploads/2020/06/An-Online-Environmental-Scan-of-Right-wing-Extremism-in-Canada-ISD.pdf>; Audette-Longo, T. (2022). The Trucker Convoy provides a case study in online threats and IRL harms. Available at: <https://j-source.ca/the-trucker-convoy-provides-a-case-study-in-online-threats-and-irl-harms/>; Awan, I., & Zempi, I. (2016). The affinity between online and offline anti-Muslim hate crime: Dynamics and impacts. *Aggression and violent behavior*, 27, 1-8.

[24] Williams, M. L., Burnap, P., Javed, A., Liu, H., & Ozalp, S. (2020). Hate in the machine: Anti-Black and anti-Muslim social media posts as predictors of offline racially and religiously aggravated crime. *The British Journal of Criminology*, 60(1), 93-117.

[25] Hoffman, B., Ware, J., & Shapiro, E. (2020). Assessing the threat of incel violence. *Studies in Conflict & Terrorism*, 43(7), 565-587.

[26] Cnockaert, J. (2022). Hate symbols at convoy protest bolster urgency for online hate legislation, say advocacy group, NDP. The Hill Times. Available at: <https://www.hilltimes.com/2022/02/14/hate-symbols-at-convoy-protest-bolster-urgency-for-online-hate-legislation-say-advocacy-group-ndp/343814>

[27] Centre on Hate, Bias and Extremism (CHBE) (2021). CHBE Statement on the Murder of Muslim Family in London, Ontario. Available at: <https://socialscienceandhumanities.ontariotechu.ca/centre-on-hate-bias-and-extremism/media/current-commentaries.php>

[28] Panneton, D. (2022) The trucker convoy shows how Canadians are being sucked into larger conspiratorial narratives. The Globe and Mail. Available at: <https://www.theglobeandmail.com/opinion/article-the-trucker-convoy-shows-how-canadians-are-being-sucked-into-larger/>

[29] Davey, J. (2020). Infiltration operations: How 4chan sought to compromise the Black Lives Matter protests. Institute for Strategic Dialogue. Available at: https://www.isdglobal.org/digital_dispatches/infiltration-operations-how-4chan-sought-to-compromise-the-black-lives-matter-protests/

[30] Perkel, C. (2021). 51% spike in hate crimes spurred partly by pandemic, Toronto police report finds. CBC News. Available at: <https://www.cbc.ca/news/canada/toronto/51-spike-in-hate-crimes-spurred-partly-by-pandemic-toronto-police-report-finds-1.5998703>

[31] Brandwatch and Ditch the Label (2021). Uncovered: Online Hate Speech in the Covid Era. Available at: https://www.ditchthelabel.org/wp-content/uploads/2021/11/Uncovered_Online_Hate_Speech_DTLxBW_V2-1.pdf

[32] The term QTBIPOC stands for Queer, Trans, Black, Indigenous, People of Color and acknowledges the particular challenges of Black, Indigenous and Racialized people within queer and trans communities.

[33] Calls for focus group participants invited youth aged 16-30 years across Canada.

[34] Environics Research and YWCA Canada (2022). Experience with Online Hate: A national survey to better understand young women and gender diverse people's experiences with online hate in Canada. Report of Findings, September 8, 2022.

[34A] Public Policy Forum. (2019). Policy Approaches to Online Hate. Summary Report. Available at: <https://ppforum.ca/wp-content/uploads/2019/06/Summary-Report-Policy-Approaches-to-Hate-Online-EN-Final.pdf>

[35] Environics Research and YWCA Canada (2022). Experience with Online Hate: A national survey to better understand young women and gender diverse people's experiences with online hate in Canada. Report of Findings, September 8, 2022.

[36] Environics Research and YWCA Canada (2022). Experience with Online Hate: A national survey to better understand young women and gender diverse people's experiences with online hate in Canada. Report of Findings, September 8, 2022.

[37] An online multiplayer game that was designed for children between 6-14 years by Disney Canada Inc.

[38] Thomson, A. (2017). Concerns raised as report suggests Canadians spending more time online. The Globe and Mail. Available at: <https://www.theglobeandmail.com/news/national/concerns-raised-as-report-suggests-canadians-spending-more-time-online/article34360751/?cmpid=rss1&click=dlvr.it>

[39] Liewicki, N. (2022). Online youth victimization up 37%, according to Canadian Centre for Child Protection. CBC. Available at: <https://www.cbc.ca/news/canada/manitoba/online-youth-victimization-canadian-centre-child-protection-1.6343186>

[40] Brisson-Boivin, K. (2019). Pushing Back Against Hate Online. MediaSmarts. Ottawa.

[41] Light (2020). Rising Levels of Hate Speech & Online Toxicity During This Time of Crisis. Available at: https://light.com/Toxicity_during_coronavirus_Report-L1ght.pdf

[42] Fisher, M. and Taub, A. (2018). How Everyday Social Media Users Become Real-World Extremists. The New York Times. Available at: <https://www.nytimes.com/2018/04/25/world/asia/facebook-extremism.html>

[43] The understanding that algorithms are not neutral and can perpetuate bias and inequality.

[44] Ghaffary, S. (2019). The algorithms that detect hate speech online are biased against black people. Vox. Available at: <https://www.vox.com/recode/2019/8/15/20806384/social-media-hate-speech-bias-black-african-american-facebook-twitter>

[45] Lim, M and Alrasheed, G. (2021). Beyond a technical bug: Biased algorithms and moderation are censoring activists on social media. The Conversation. Available at: <https://theconversation.com/beyond-a-technical-bug-biased-algorithms-and-moderation-are-censoring-activists-on-social-media-160669>

[46] United Nations Educational, Scientific and Cultural Organization (UNESCO) (2021). Addressing hate speech on social media: contemporary challenges. Discussion Paper 01. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000379177?mkt_tok=Njl0LVhFTC01NTIAAAGGgz6babLAW4ZBNLjn21ltHYBOutawOZtFx5oO1btu6LMMnnFw_2i-jq_kZ_FUPJnmGFL7xWN29PI70XspDA

[47] Office of the High Commissioner for Human Rights (OHCHR) (2021). Moderating online content: fighting harm or silencing dissent? Available at: <https://www.ohchr.org/en/stories/2021/07/moderating-online-content-fighting-harm-or-silencing-dissent>

[48] See the following articles for discussion on how proponents for fewer restrictions on freedom of expression frame government anti-hate interventions as excessively restricting individual rights. Scharfstein, L. E. (2019). The Hate Speech Debate: The Supreme Court, the Federal Government, and the Need for Civil Hate Speech Provisions. *Asper Review of International Business and Trade Law*, 19, 375 and Walker, J. (2018). Hate speech and freedom of expression: Legal boundaries in Canada. Background Paper. Library of Parliament, Ottawa.

[49] Khoo, C. (2021). Deplatforming Misogyny. Report on Platform Liability for Technology-Facilitated Gender-Based Violence. Women's Legal Education and Action Fund (LEAF). Toronto, Ontario. Available at: <https://www.leaf.ca/publication/deplatforming-misogyny/>

[50] Khoo, C. (2021). Deplatforming Misogyny. Report on Platform Liability for Technology-Facilitated Gender-Based Violence. Women's Legal Education and Action Fund (LEAF). Toronto, Ontario. Available at: <https://www.leaf.ca/publication/deplatforming-misogyny/>

[51] Plan International. (2020). Free to be Online? Girls' and Young women's Experiences of Online Harassment. The State of the World's Girls Report. Available at: <https://plan-international.org/publications/free-to-be-online/>

[52] Center for Countering Digital Hate (CCDH). (2022). Hidden Hate: How Instagram fails to act on 9 in 10 reports of misogyny in DMs. Available at: <https://counterhate.com/wp-content/uploads/2022/05/Final-Hidden-Hate.pdf>

[53] United Nations Educational, Scientific and Cultural Organization (UNESCO). (2015) Cyber violence against women and girls. A world-wide wake-up call. Defining the threat environment: the 'cyber' nature of VAWG. Available at: <https://en.unesco.org/sites/default/files/gender-report2015final.pdf>

[54] Canadian Race Relations Foundation (CRRF). (2021). Standing up to hate will take collective action. Available at: <https://www.crrf-fcrr.ca/en/news-a-events/articles/item/27495-standing-up-to-hate-will-take-collective-action>

[55] Salovarra, J. (2022). Online Harassment Is Not Gender-Neutral. UN Chronicle. Available at: <https://www.un.org/en/un-chronicle/online-harassment-not-gender-neutral>

[56] Environics Research and YWCA Canada (2022). Experience with Online Hate: A national survey to better understand young women and gender diverse people's experiences with online hate in Canada. Report of Findings, September 8, 2022.

[57] Online hate against women and gender diverse public personalities is designed to disparage, defame, discredit and endanger women and gender diverse people with a view to 'put them in their place' and discourage or 'chill their active participation' in public life. See: Posetti, J. et al. (2021). The Chilling: Global trends in online violence against women journalists. Research Discussion Paper. UNESCO. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000377223/PDF/377223eng.pdf.multi>

[58] Mijatović, D. (2022). No space for violence against women and girls in the digital world. Council of Europe, Commissioner for Human Rights. Human Rights Comments. Available at: <https://www.coe.int/en/web/commissioner/-/no-space-for-violence-against-women-and-girls-in-the-digital-world>

[59] UN Women– Asia Pacific. (2021). Eliminating Online Hate Speech to Secure Women's Political Participation. Action Brief. Available at: <https://asiapacific.unwomen.org/en/digital-library/publications/2021/04/eliminating-online-hate-speech-to-secure-women-s-political-participation>

[60] UN Human Rights. (2017). UN experts urge States and companies to address online gender-based abuse but warn against censorship. Joint statement by UN Special Rapporteur on the Promotion of Freedom of Expression and UN Special Rapporteur on Violence Against Women. Available at: <https://www.ohchr.org/en/press-releases/2017/03/un-experts-urge-states-and-companies-address-online-gender-based-abuse-warn>

[61] A subset of online harassment involving the coordinated creation and/or circulation of deceptive or inaccurate information and images against people in politics or public life based on their gender and misplaced gender and sex-based narratives against women and gender diverse people. See more: Jankowicz, N., Hunchak, J., & Pavluic, A. (2021). Malign Creativity: How gender, sex and lies are weaponized against women online. Wilson Centre. Available at: <https://www.wilsoncenter.org/publication/malign-creativity-how-gender-sex-and-lies-are-weaponized-against-women-online>

[62] For non-exhaustive coverage of this issue see: The Canadian Press. (2017). 'I became a target': The difficult tenure of women politicians in Canada. CBC. Available at: <https://www.cbc.ca/news/canada/newfoundland-labrador/women-politics-canada-1.4231672> and Pulfer, R. (2022). Opinion: The online hate faced by women in journalism is an issue that affects us all. The Globe and Mail. Available at: <https://www.theglobeandmail.com/opinion/article-the-online-hate-faced-by-women-in-journalism-is-an-issue-that-affects/>

[63] Plan International. (2020). Free to be Online? Girls' and Young women's Experiences of Online Harassment. The State of the World's Girls Report. Available at: <https://plan-international.org/publications/free-to-be-online/>

[64] GamerGate was an online hate and harassment campaign that targeted women, feminists and diversity in gaming culture and the video game industry. This backlash led to one of the first prominent disengagements of women from online discourse.

[65] Publishing an individual's private or identifying information online, especially as a form of punishment, revenge, extortion, or shaming. See Liewicki, N. (2022). Online youth victimization up 37%, according to Canadian Centre for Child Protection. CBC. Available at: <https://www.cbc.ca/news/canada/manitoba/online-youth-victimization-canadian-centre-child-protection-1.6343186>

[66] A practice involving someone calling in a threat to police, resulting in armed officers being sent to another person's home or work. See Dubinski, K. (2022). Trans Twitch star arrested at gunpoint fears for life after someone sent police to her London, Ont., home. CBC. Available at: <https://www.cbc.ca/news/canada/london/trans-twitch-star-arrested-at-gunpoint-fears-for-life-after-someone-sent-police-to-her-london-ont-home-1.6546015>

[67] UN Women. (2022) Accelerating Efforts to Tackle Online and Technology Facilitated Violence Against Women and Girls (VAWG). Available at: https://www.unwomen.org/sites/default/files/2022-10/Accelerating-efforts-to-tackle-online-and-technology-facilitated-violence-against-women-and-girls-en_0.pdf and Posetti, J. et al. (2021). The Chilling: Global trends in online violence against women journalists. Research Discussion Paper. UNESCO. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000377223/PDF/377223eng.pdf.multi>

[68] Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). Countering Online Hate Speech. Paris: UNESCO.

[69] YWCA Canada. (2019). Addressing Online Hate: Applying and Intersectional Gender Lens. Brief submitted to Standing Committee on Justice and Human Rights. Available at: <https://www.ourcommons.ca/Content/Committee/421/JUST/Brief/BR10528526/br-external/YWCACanada-e.pdf>

[70] Orlando, J. (2020). Young People are exposed to more hate online during COVID. And it risks their health. The Conversation. Available at: <https://theconversation.com/young-people-are-exposed-to-more-hate-online-during-covid-and-it-risks-their-health-148107>

[71] Kid's Help Phone. (2022). Helping Young People Recognize and Respond to Online Hate. Available at: <https://kidshelpphone.ca/get-info/helping-young-people-recognize-and-respond-to-online-hate>

[72] Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). Countering Online Hate Speech. Paris: UNESCO.

[73] MediaSmarts. (2019). Online Hate: Responses and Solutions in the Classroom. Available at: <https://mediasmarts.ca/online-hate/responses-solutions>

[74] Council of Agencies Serving South Asians (CASSA). (2022). Eradicate Hate Toolkit. A Toolkit for Combating Online Hate. First Edition. Available at: <https://static1.squarespace.com/static/5e09162ecf041b5662cf6fc4/t/620588acf4c1171b5c847300/1644529840547/English-Eradicate-Hate-Toolkit.pdf>

[75] Council of Europe. (2017). No Hate Speech Youth Campaign. Available at: <https://www.coe.int/en/web/no-hate-campaign/no-hate-speech-movement>

[76] Blank, A. (2019). Combatting Online Hate: An Alternative Approach. Brief submitted to Standing Committee on Justice and Human Rights. Available at: <https://www.ourcommons.ca/Content/Committee/421/JUST/Brief/BR10510667/br-external/BlankAlyssa-e.pdf>

[77] Gilbert + Tobin. (2022). New Approaches to Combatting Online Hate. Digital Hub. Available at: <https://www.gtlaw.com.au/knowledge/new-approaches-combatting-online-hate-speech>

[78] Gilbert + Tobin. (2022). New Approaches to Combatting Online Hate. Digital Hub. Available at: <https://www.gtlaw.com.au/knowledge/new-approaches-combatting-online-hate-speech>

[79] See O'Callaghan, D., Greene, D., Conway, M., Carthy, J., & Cunningham, P. (2015). Down the (white) rabbit hole: The extreme right and online recommender systems. *Social Science Computer Review*, 33(4), 459-478 and Khan, A. R. (2018). How the internet may be turning us all into radicals. Maclean's. Available at: <https://www.macleans.ca/society/technology/how-the-internet-may-be-turning-us-all-into-radicals/> quoted in Public Policy Forum. (2019). Policy Approaches to Online Hate. Summary Report. Available at: <https://ppforum.ca/wp-content/uploads/2019/06/Summary-Report-Policy-Approaches-to-Hate-Online-EN-Final.pdf>

[80] A racist, xenophobic, white nationalist far-right conspiracy theory that believes that policies such as immigration policies are designed to systematically undermine and 'replace' the culture, social-economic standing, and political power of white people of European descent living in Western countries.

[81] A forum or message board dedicated to a specific topic on the website Reddit.

[82] A software application that runs automated tasks on the Internet, usually with the intent to imitate online human activity. They can be programmed to perform functions such as mass messaging at high speeds and scale.

[83] This is especially pertinent considering recent news of the departure of Meta's Global Manager of Community Development after allegedly engaging in graphic and inappropriate online communications with a minor. Canales, K. (2022). Meta says a manager is no longer at the company after allegedly appearing in video from a civilian underage sex sting group. *Business Insider*. Available at: <https://www.businessinsider.com/meta-facebook-manager-departs-company-following-video-sting-operation-2022-2>

[84] Meta (2021). Community Standards Enforcement Report, Third Quarter 2021. Available at: <https://about.fb.com/news/2021/11/community-standards-enforcement-report-q3-2021/>

[85] Participants in two focus groups raised the idea of empathetic and non-aggressive responses to "meet hate with kindness". Many mentioned witty one-liners and snappy comebacks they used to deflect hate speech and demonstrate alternative approaches to the speaker. Among their suggestions was a 'kind comeback generator', a technological tool that utilizes Artificial Intelligence to offer positive and productive responses that are visible alongside or after a post is removed. See: Gilbert + Tobin. (2022). *New Approaches to Combatting Online Hate*. Digital Hub. Available at: <https://www.gtlaw.com.au/knowledge/new-approaches-combatting-online-hate-speech>

[86] Sultan, A. (2019). We Need Real Transparency About Hate on Social Media. ADL Center on Technology and Society. Available at: <https://www.adl.org/blog/we-need-real-transparency-about-hate-on-social-media>

[87] Public Policy Forum. (2019). Policy Approaches to Online Hate. Summary Report. Available at: <https://ppforum.ca/wp-content/uploads/2019/06/Summary-Report-Policy-Approaches-to-Hate-Online-EN-Final.pdf>

[88] Walker, J. (2018). Hate speech and freedom of expression: Legal boundaries in Canada. Background Paper. Library of Parliament, Ottawa. Available at: https://lop.parl.ca/sites/PublicWebsite/default/en_CA/ResearchPublications/201825E#txt12

[89] MediaSmarts. (2019). Online Hate and Canadian Law. Available at: <https://mediasmarts.ca/digital-media-literacy/digital-issues/online-hate/online-hate-canadian-law>

[90] Public Policy Forum. (2019). Policy Approaches to Online Hate. Summary Report. Available at: <https://ppforum.ca/wp-content/uploads/2019/06/Summary-Report-Policy-Approaches-to-Hate-Online-EN-Final.pdf>

[91] Canadian Race Relations Foundation (CRRF). (2021). New Study gives voice to Canadians on racism, discrimination and action. National Survey conducted by Nanos for the CRRF, November 2021. Available at: https://www.crrf-fcrr.ca/images/100_Day_Poll_2021/2021-1995_CRRF_-_Populated_report_-_FINAL_with_tabs6.pdf

[92] Connolly, A. (2022). Trudeau promised online hate reform by Feb. 3. Experts say it's unlikely, but urge action – National. *Global News*. Available at: <https://globalnews.ca/news/8586520/online-hate-legislation-canada/>

Funded by the Government of Canada
Financé par le gouvernement du Canada

Canada



Public Safety
Canada

Sécurité publique
Canada

ywcacanada.ca

YWCA Canada gratefully acknowledges Anishinaabe Onyota'a:aka artist Tsista Kennedy and The Indigenous Friends Association, for the artwork commissioned and featured in this report.