

Grand Moot 2022

Freedom of Expression and the Regulation of Social Media

R v ChitChat

Official Problem by Charles Ma and Mackenzie Faulkner

R v ChitChat

1. This appeal addresses the constitutionality of a law which prohibits viewpoint-based discrimination on privately held social media platforms by limiting said platforms' ability to censor certain content. It explores the balance between addressing misinformation and offensive speech while ensuring that public discourse is not unduly influenced by the decisions of private corporations.
2. The appeal takes place in Falconer, a common law province in the country of Flavelle. The Constitution, judicial system, statutory law, common law, and social and political history of Flavelle and Falconer are identical to those of Canada and Ontario, respectively.
3. Flavelle's highest court is the Supreme Court of Flavelle. All Canadian legislation is binding on the Supreme Court of Flavelle, but the Court is not bound by Canadian jurisprudence. However, decisions of Canadian courts, particularly the Supreme Court of Canada, are considered highly persuasive.
4. The Superior Court of Falconer, the Falconer Court of Appeal, and the Supreme Court of Flavelle all have jurisdiction over the issues raised below.

Facts

5. The following information was gathered throughout the course of ChitChat's *Charter* application before the Superior Court of Falconer. It represents the totality of the evidence deemed relevant to the ss. 2(b) and 15 applications.

ChitChat

6. ChitChat is a social media platform created by and for Flavellians. The platform is incredibly popular, boasting 9.5 million monthly active users in a nation of 38 million. Its content is entirely user-generated, largely consisting of discussion about popular culture and current affairs. Registered users may publish posts of up to 250 characters. They may also publish pictures and videos. ChitChat users collectively create 7,000 posts per minute, totalling more than 10 million posts per day.

7. ChitChat does not generate revenue directly from user content. Nearly the entirety of its revenue comes from placing advertisements across the platform, paid for by corporations or individuals.

8. Although ChitChat is not the only social media platform operating in Flavelle, the vast majority of the market for social media is controlled by ChitChat and a small number of other dominant platforms, all of which use similar Terms of Service.

9. To register for ChitChat, users need only provide an email address. Registration enables users to create a personalized profile, publish their own posts, and “follow” other users. It also allows them to “block” other users; if one user blocks another, neither of the two users will be able to follow the other or see the other’s posts. People who are not registered to use ChitChat can view users’ posts but cannot publish their own posts or follow other users.

10. ChitChat requires that all users agree to its Terms of Service (“Terms”) before completing the registration process. Under its Terms, ChitChat reserves the right to take various enforcement actions if it determines that a user’s content is “harmful or abusive.” The Terms do not provide any further specificity; ChitChat deliberately worded this policy broadly in order to retain broad discretion over the limits of acceptable content on the platform.

11. When a user’s content is deemed to be harmful or abusive, ChitChat may:

- Delete the post;
- Suspend or ban the user;
- Limit the visibility of the user or the post (colloquially known as “shadow banning”); or
- Add labels or overlays to the post in order to identify it as containing false or misleading information.

12. To enforce its Terms, ChitChat relies on a combination of proactive and reactive content moderation practices. ChitChat’s proprietary technology proactively censors content containing certain words and phrases. ChitChat also reactively moderates content by allowing users to report content for ChitChat’s review. ChitChat then determines whether the reported content is “harmful or abusive,” thus warranting further action. These determinations are made at the sole discretion

of ChitChat’s compliance team. In recent years, ChitChat has experimented with outsourcing this discretion to external corporate partners.

13. Users who disagree with any particular exercise of this discretion may appeal the decision. Appeals are reviewed by senior employees of the compliance staff, who either affirm or alter the original sanction. There is no further appeals process once this second decision has been rendered.

14. Given its sheer size, ChitChat’s content moderation practices remain imperfect; some harmful posts remain unaddressed while some innocuous posts are removed. Nonetheless, ChitChat’s Founder and Chief Executive Officer, Charles Mackenzie, remains committed to making the platform a safe environment for users, free from offensive speech and misinformation.

The Digital Public Squares Act

15. In September 2021, a new majority government was elected following Flavelle’s federal election. Members of the newly formed government sought to address a concern raised by voters on the campaign trail: social media censorship. Voters complained about having their posts deleted, being suspended from platforms, or being banned based on their expressed political opinions. According to estimates calculated by reputable polling organizations, at least half of Flavellians believe that social media has hurt, rather than encouraged, open debate. The new government agreed that in a free and democratic society, such immense power over public discourse should not rest with privately held corporations.

16. To circumscribe the influence of social media platforms, the Government of Flavelle enacted the *Digital Public Squares Act* (“DPSA”). The full text of the *DPSA* is reproduced in its entirety in Appendix B. The relevant statutory text is excerpted below:

Prohibition on Censorship

3 A social media platform may not censor a user, a user’s expression, or a user’s ability to receive the expression of another person based on:

- (a) the viewpoint of the user or another person; or
- (b) the viewpoint represented in the user’s expression.

Exceptions

4(1) This Act does not prohibit a social media platform from censoring expression that:

- (a) directly incites criminal activity or consists of specific threats of violence targeted against a person or group;
- (b) depicts sexual exploitation or physical or sexual abuse; or
- (c) is otherwise unlawful.

(2) A notice which states that a user's expression might have been censored but for the provisions of this Act does not itself constitute censorship.

[...]

Sanction

6 Every one who fails to comply with section 3 of this Act is guilty of an offence and on conviction is liable for a fine of \$500.00 for each violation.

ChitChat Contravenes the DPSA

17. Following its enactment, most platforms adhered to the *DPSA*. Some of them began displaying a small text box under potentially objectionable posts, as permitted under s. 4(2) of the *DPSA*, stating: "This content may violate our Terms of Service. It is being hosted pursuant to our obligations under the *Digital Public Squares Act*."

18. Charles Mackenzie vehemently opposed the *DPSA* regime. He founded ChitChat as an open forum for civil discourse and viewed ChitChat's discretion over the platform's contents as a crucial bulwark against offensive speech and misinformation. After consulting with ChitChat's legal team, he learned that the *DPSA* may be vulnerable to a constitutional challenge. Although Mr. Mackenzie instructed his compliance staff to err on the side of caution and allow most posts to remain published, he also directed them to continue deleting any particularly egregious posts they discovered.

19. Over the next 24 hours, ChitChat deleted 1,432 posts. 323 of those posts fell within the exceptions enumerated in the *DPSA*, making their censorship lawful. The remaining 1,109 posts did not fall within any exceptions; these include, but are not limited to:

- Posts disparaging members of minority groups on the basis of their race, religion, sexual orientation, and gender identity;
- Posts asserting that COVID-19 vaccines are unsafe or ineffective;
- Posts expressing strong disapproval of Pride Month;
- Posts about a controversial Supreme Court of Flavelle ruling, some of which insinuated that the Justices in the majority were bribed or acted at the behest of politicians;
- Posts comparing the Premier of Falconer to various historical dictators after she extended the province's mask mandate;
- A post from an animal rights group which included graphic imagery of animals in factory farms;
- A post from a well-known political commentator expressing that anyone who did not vote for the incumbent government "hates their country" and that they "will spend the rest of their lives, and beyond, paying for their betrayal";
- The following post from an anonymous user with 230 followers: "Big Tech CEOs have proven themselves to be soulless oligarchs. They have continually manipulated us with zero accountability. Return the power to the people!"

20. ChitChat was charged with 1,109 violations of s. 3 of the *DPSA*. In response, ChitChat challenged the constitutionality of the *DPSA* on two grounds. First, it argued that the *DPSA* infringes s. 2(b) of the *Charter*, which guarantees freedom of expression, by compelling ChitChat to host expression against its wishes. Second, it argued that the *DPSA* infringes s. 15 of the *Charter*, the right to equality, due to the increase in offensive and derogatory expression on the platform since the *DPSA*'s enactment, which disproportionately affects members of the LGBTQ+ community and racial and religious minority groups. ChitChat argued that neither infringement is justified under s. 1 of the *Charter*.

21. The Attorney General of Flavelle has conceded that ChitChat has standing to challenge the *DPSA* on both grounds.¹ However, it argues that the *DPSA* does not infringe s. 2(b) or s. 15 and that, even if it does infringe one or both sections, it is justified under s. 1.

22. Pending the disposition of its *Charter* challenge, ChitChat has agreed not to delete any more posts apart from those deemed to clearly fall within one of the *DPSA*'s exceptions. The Attorney General of Flavelle has taken no issue with ChitChat's conduct since.

Procedural History

The Decision of the Superior Court of Falconer

23. The application judge, Salamat J, found that the *DPSA* infringes both s. 2(b) and s. 15 of the *Charter* in a manner that cannot be justified under s. 1.

24. On the s. 2(b) issue, Salamat J accepted ChitChat's submission that the effect of the *DPSA* is to compel the platforms' expression. She wrote in part:

Flavellian jurisprudence has long held that freedom of expression entails both the right to say something and the right to say nothing. Where a person or corporation is forced by the state to speak despite wishing to remain silent, that freedom is infringed. It is no different when a person or corporation is forced by the state to host or facilitate expression with which it disagrees. Stripping social media platforms of the ability to remove certain expression negates their freedom to only associate themselves with the expression which they wish to associate with. This necessarily infringes s. 2(b) of the *Charter*.

25. On the s. 15 issue, Salamat J granted leave to hear evidence from two witnesses in order to fully understand the adverse effects being claimed by ChitChat.

26. First, ChitChat's Director of Compliance, Emily Hean, testified that she was unsure whether there had been an increase in the volume of offensive content disparaging minority groups being posted since the *DPSA*'s enactment. However, she testified that, over the last three years, ChitChat's compliance staff had deleted an average of 110,000 posts per day (approximately 1%

¹ *R v Big M Drug Mart Ltd.*, [1985] 1 SCR 295 at para 39: "Any accused, whether corporate or individual, may defend a criminal charge by arguing that the law under which the charge is brought is constitutionally invalid."

of total daily posts) for violating the policy prohibiting “harmful and abusive” content. Much of this deleted content included slurs describing members of minority groups and invoked prejudicial stereotypes. Based on this evidence, ChitChat argued that, while there may not be an increased level of such content being posted, the *DPSA* markedly increases the number of harmful and abusive posts to which users are subjected since the platforms can no longer delete these posts.

27. Next, the Court heard from Dr. Kathryn Mullins, a sociology professor at the University of Falconer who has conducted extensive research on the effects of offensive expression. She testified that, based on her research, roughly 70% of such expression is targeted at members of minority groups. She further noted that when people face greater exposure to, or are directly or indirectly targeted by, this kind of expression, they may experience heightened stress, anxiety, depression, increased drug and alcohol use, lower self-esteem, and other psychological symptoms such as pain, fear, and intrusive thoughts of intimidation and denigration.

28. Accordingly, Salamat J found that, although the law is facially neutral, its harmful effects are profound and disproportionately borne by members of minority groups:

The effects of the *DPSA* are not experienced uniformly across the Flavellian population. Members of certain groups—namely, members of minority groups—bear the burden of unfiltered and unmoderated expression more than others. It is undisputed that even though some of this expression is bigoted and wrongheaded, it is nonetheless lawful expression, rendering it ineligible for valid censorship under s. 4(1)(c) of the *DPSA*. However, its legality does not make it any less harmful. Exposure to, and victimization from, such expression can entail serious ramifications. These ramifications, which are disproportionately and acutely experienced by members of the LGBTQ+ community and religious and racial minority groups, are a direct result of the *DPSA*. Consequently, I find that the *DPSA* infringes s. 15 of the *Charter*.

29. Finally, Salamat J ruled that these infringements could not be justified under s. 1. She held that encouraging viewpoint diversity and facilitating robust public discourse was a pressing and substantial purpose. She also found that there was a rational connection between that purpose and the *DPSA*’s prohibition of viewpoint-based censorship. However, she found that the law was not minimally impairing. She wrote:

While the *DPSA*’s scope is limited to “social media platforms,” a carefully defined term in the statute, it effectively prohibits platforms from doing anything about offensive speech and misinformation on their platforms at all. A blanket

prohibition on censoring *any* expression based on viewpoint (apart from permitting mere acknowledgements of the *DPSA*) precludes any good-faith case-by-case approach to social media expression. As a result, I cannot find that these infringements are minimally impairing. They cannot be justified under s. 1.

30. The *DPSA* was declared of no force or effect, and ChitChat's charge was quashed. The Attorney General of Flavelle appealed the decision to the Falconer Court of Appeal.

The Decision of the Falconer Court of Appeal

31. The majority of the Falconer Court of Appeal allowed the appeal. Each of the three justices wrote a separate opinion.

32. Grondin JA held that neither s. 2(b) nor s. 15 were infringed. In rejecting ChitChat's s. 2(b) claim, he wrote:

The application judge erred in assuming that compelling one's expression is the same as compelling one to host someone else's expression. This is a crucial distinction. *Lavigne v OPSEU* points to two factors which the application judge omitted: 1) whether the public would reasonably think that the message is shared or supported by the claimant, and 2) whether the claimant has an opportunity to disavow the expression.² Section 4(2) of the *DPSA* addresses both factors by allowing the platform to explain to the public the reason that it is hosting the message—namely, because of its obligations under the *DPSA*. That serves as an opportunity for the platform to distance itself from the expression and, in so doing, ensures that the public does not associate the platform with the expression. This option is available to all social media platforms and most are already making use of it. Therefore, I cannot conclude that the *DPSA* compels the platforms' expression.

33. In dismissing the s. 15 claim, Grondin JA wrote:

I cannot accept the claimant's s. 15 argument for two reasons. First, the government is not constitutionally obligated to protect its citizens from the opinions of their peers, no matter how odious those views may be. Indeed, the antidote to harmful expression is more expression, not less. Shining light on these divergent opinions, which exist irrespective of the *DPSA*, cannot constitute an infringement of the right to equality. Second, even if legally cognizable adverse effects do exist, they are not an unavoidable aspect of the lives of those who are affected by the expression at issue. Using social media is neither essential nor

² *Lavigne v Ontario Public Service Employees Union*, [1991] 2 SCR 211 at 278-279.

fundamental to one's life; the choice to log off and remove oneself from the situation always remains. In other words, using social media is merely a lifestyle choice. This cuts against the adversity of the effects being claimed.

34. Park JA concurred with Grondin JA's finding that the *DPSA* was constitutional, but did so for considerably different reasons. Park JA endorsed the lower court's holdings on ss. 2(b) and 15, holding that both sections were infringed. However, he held the infringements to be justified under s. 1. He agreed with the lower court that there was a pressing and substantial purpose, and that the *DPSA*'s measures were rationally connected to that purpose. With respect to minimal impairment, he held that the law fell within a reasonable range of alternatives and deferred to Parliament's policy decision. For the final stage of the s. 1 analysis, the balancing of the salutary and deleterious effects, he wrote:

I cannot conclude that the *DPSA* is an unjustifiable infringement of ss. 2(b) and 15. To do so would set a troubling precedent, significantly constraining future legislative decisions. My greatest concern lies with the claimant's s. 15 argument. A law which facilitates more expression will invariably also facilitate some harmful expression. If this law is held to be an unjustifiable infringement of s. 15 merely because it facilitates harmful expression, then so would any law which provides general protections for expression. This is a critical factor which favours the *DPSA* being justified; I have not been referred to any countervailing factor that outweighs it. Accordingly, striking down this law is inappropriate. I find that the *DPSA* is justified under s. 1.

35. Rand JA dissented, adopting the reasons of the lower court in their entirety. In response to Grondin JA's contention that social media is a "lifestyle choice," she wrote:

As the Supreme Court of Canada has acknowledged, "[h]aving the choice to remain "offline" may not be a real choice in the Internet era."³ Social media platforms now act as essential forums for debate and discourse on all matters of social and political importance. Today's politicians even use social media to make official announcements. For members of minority groups, the price of civic participation cannot be enduring bigotry. Functionally, the *DPSA* requires just that.

³ *Douez v Facebook, Inc*, 2017 SCC 33 at para 56.

36. Since the majority of the Court found the *DPSA* to be constitutionally valid, the charges against ChitChat were restored.

Issues on Appeal

37. ChitChat has appealed the Falconer Court of Appeal's decision as of right to the Supreme Court of Flavelle. The Court is being asked to decide the following issues:

1. Does the *DPSA* infringe s. 2(b) of the *Charter*?
2. Does the *DPSA* infringe s. 15 of the *Charter*?
3. If the *DPSA* infringes either s. 2(b) or s. 15, is/are the infringement(s) justified under s.1 of the *Charter*?

**APPENDIX A: RELEVANT PROVISIONS OF THE *FLAVELLIAN CHARTER OF
RIGHTS AND FREEDOMS***

1 The *Flavellian Charter of Rights and Freedoms* guarantees the rights and freedoms set out in it subject only to such reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society.

2 Everyone has the following fundamental freedoms:

(b) freedom of thought, belief, opinion and expression, including freedom of the press and other media of communication;

15(1) Every individual is equal before and under the law and has the right to the equal protection and equal benefit of the law without discrimination and, in particular, without discrimination based on race, national or ethnic origin, colour, religion, sex, age or mental or physical disability.

(2) Subsection (1) does not preclude any law, program or activity that has as its object the amelioration of conditions of disadvantaged individuals or groups including those that are disadvantaged because of race, national or ethnic origin, colour, religion, sex, age or mental or physical disability.

APPENDIX B: THE *DIGITAL PUBLIC SQUARES ACT*

Purpose of this Act

1(1) The purpose of this Act is to:

- (a) affirm that free and open public discourse is critical to the functioning of a free and democratic society;
- (b) recognize that social media platforms play an increasingly central role in facilitating and hosting public discourse on matters of social and political importance; and
- (c) promote the marketplace of ideas, thereby facilitating the pursuit of truth, individual self-fulfillment, and democratic participation.

(2) In furtherance of that purpose, this Act creates measures intended to reduce or eliminate discrimination on the basis of viewpoint on social media platforms.

Interpretation

2 In this Act,

social media platform means an Internet website or application that is open to the public, allows a user to create an account, and enables users to communicate with other users for the primary purpose of posting information, comments, messages, or images. The term does not include:

- (a) an Internet service provider;
- (b) electronic mail; or
- (c) an online service, application, or website:
 - (i) that consists primarily of news, sports, entertainment, or other information or content that is not user-generated but is preselected by the provider; and
 - (ii) for which any chat, comments, or interactive functionality is incidental to, directly related to, or dependent on the provision of the content described in subparagraph (i);

censor means any action taken to edit, alter, block, ban, delete, deplatform, demonetize, regulate, remove, restrict, inhibit the publication or reproduction of, deny equal access or visibility to, or suspend a right to post.

receive, with respect to an expression, means to read, hear, look at, access, or gain access to the expression;

user means a person who posts, uploads, transmits, shares, or otherwise publishes or receives expression, through a social media platform. The term includes a person who has a social media platform account that the social media platform has disabled or locked.

Prohibition on Censorship

3 A social media platform may not censor a user, a user's expression, or a user's ability to receive the expression of another person based on:

- (a) the viewpoint of the user or another person; or
- (b) the viewpoint represented in the user's expression.

Exceptions

4(1) This Act does not prohibit a social media platform from censoring expression that:

- (a) directly incites criminal activity or consists of specific threats of violence targeted against a person or group;
- (b) depicts sexual exploitation or physical or sexual abuse; or
- (c) is otherwise unlawful.

(2) A notice stating that a user's expression might have been censored but for the provisions of this Act does not itself constitute censorship.

Waiver

5 A waiver or purported waiver of the protections provided by this Act is void as unlawful and against public policy, and a court or arbitrator may not enforce or give effect to the waiver, notwithstanding any contractual choice-of-law provisions.

Sanction

6 Every one who fails to comply with section 3 of this Act is guilty of an offence and on conviction is liable for a fine of \$500.00 for each violation.