

Torrential Twitter: Climate Change, Female Politicians, and Harassment - Codebook*

Inessa De Angelis

July 12, 2023

Table of contents

1	Overview	1
2	Coding Schedule	2
2.1	Severity of Harassment	2
2.2	Type of Account	5
3	Other Guidelines	6
3.1	Retweets	6
3.2	Threaded Responses	6
3.3	Media	6
3.4	Hate Speech	7
	References	7

1 Overview

This codebook is designed to inform the coding process for my Torrential Twitter trial study. Prior to coding Tweets, the raw data sets were initially cleaned using the statistical programming software R (R Core Team 2023), utilizing functions from **Janitor** (Firke 2023). Tweets will then be coded using an application designed with Shiny (Chang et al. 2023) to determine the severity of harassment and type of account.

Severity of Harassment includes:

*Code and other information available at: https://github.com/InessaDeAngelis/Torrential_Twitter_Trial

- Positive
- Neutral
- Name-calling
- Gender-based insults
- Vicious language
- Credible threats
- Hate speech

Type of account includes:

- Personal
- Professional
- Bots
- Spammers
- Anonymous
- Suspended/Deleted

2 Coding Schedule

2.1 Severity of Harassment

The coding of sentiment in Tweets will be conducted to determine the severity of harassment, measuring Tweets on a seven-point scale ranging from positive to hate speech (Mantilla 2013), (Nadim and Fladmoe 2021). The sentiment categories will be mutually exclusive, therefore only one of the seven types of harassment can be selected per reply (Krippendorff 2011). The seven-point scale contains the following categories, and they are:

1. Positive

- a. Tweets that employ positive wording and tone without expressing complaints (Zavattaro, French, and Mohanty 2015).
- b. Punctuation, such as exclamation marks, may be used to convey a positive message (Jansen et al. 2009). See Table 1 for an example and note the use of the exclamation mark.

Table 1: Sample of a **positive** Tweet sent to MP Elizabeth May.

Tweet
@ ElizabethMay @ paulmanly The forefront of the climate revolution. They’re pushing big on Vancouver Island and for good reason!

2. Neutral

- A neutral Tweet states a matter-of-fact without expressing an opinion or exuding judgement (Jansen et al. 2009).
- No specific punctuation is used and the overall tone of the Tweet is neither positive or harassing. Reference Table 2 for an example of a user stating something that occurred during proceedings in the House of Commons.

Table 2: Sample of a **neutral** Tweet sent to MP Laurel Collins.

Tweet
@ Laurel_BC The gov’t declared a climate emergency.

3. Name-calling

- The use of derogatory and offensive language, often grounded in stereotypes to refer to female politicians, their policies, and assumed qualifications (Nadim and Fladmoe 2021), (Theocharis et al. 2020). For an example, see Table 3, where a Twitter user utilizes the term “climate barbie”.

Table 3: Sample of a Tweet containing **Name-calling** sent to former MP Catherine McKenna.

Tweet
@ cathmckenna CO2 is NOT pollution climate barbie, no matter how many times or how loudly you proclaim it

4. Gender-based insults

- Derogatory comments and terms used against women, including “whore” and “slut” which are “... designed to insult and humiliate women, especially in regard to their weight and physical appearance” (Mantilla 2013, 564). See Table 4 for an example where the user references MP McKenna’s physical appearance.

Table 4: Sample of a Tweet containing **Gender-based insults** sent to former MP Catherine McKenna.

Tweet
@ cathmckenna Hey Blonde Bimbo Barbie why do you look like you always just scraped outta bed

5. Vicious language

- a. Descriptions using vile language of behaviour that gendertrolls (who are men) claim that they would like to subject the female user to (Mantilla 2013). Table 5 provides an example.

Table 5: Sample of Tweet containing **Vicious Language** sent to former MP Catherine McKenna (language redacted for this document).

Tweet
@ cathmckenna @ carleton_ F\$ck you climate b*tch

6. Credible Threats

- a. Tangible threats of violent sexual behaviour, rape, doxxing, and stalking which the gendertrolls would like to subject the female user to (Mantilla 2013). This includes seeking out the female politician with the intent to cause harm (Vickery and Everbach 2018). Reference Table 6 for an example.

Table 6: Sample of Tweet containing **Credible Threats** sent to former MP Catherine McKenna.

Tweet
@ cathmckenna U DO KNOW THAT CANADIANS ARE TOTALLY WAKING UP TO YOUR LIES AND AFTER YOU LOSE THE ELECTION WE WILL MOST DEFINITELY HUNT YOU DOWN AND HAVE YOU ARRESTED AND TRIED FOR TREASON WITH YOUR DISGUSTING CARBON AND CLIMATE LIES U WILL BE DEALT WITH !!!!!

7. Hate speech

- a. Hateful comments which oppose freedom of speech, violate human rights, and may even be unlawful (Nadim and Fladmoe 2021) (Chetty and Alathur 2018)
- b. Online or offline vilification of an individual communicated in a specific context based on “... their immutable characteristics, such as their race, ethnicity, national origin, religion, gender, gender identity, sexual orientation, age, or disability” (Carlson 2021, 4).

- c. Any form of expression that “... seeks to promote, or has the capacity to increase hatred against a person or group of people because of a characteristic they share, or a group to which they belong” (Saleem et al. 2017, 1)
- d. See Table 7 for an example of the most hateful speech found in this sample and the [Other Guidelines](#) section for notes on the frequency of hate speech found in this trial.

Table 7: Sample of Tweet containing **Hate Speech** sent to former MP Catherine McKenna (language redacted for this document).

Tweet
@ cathmckenna Better polish up that resume, climate mckenn*zi.

2.2 Type of Account

Only one type of account can be selected and will also be coded mutually exclusively. If one user replies to a single Tweet by a female politician using multiple separate Tweets, then each Tweet will be coded separately. There will be no limit on the coding of the number of replies sent by an individual account.

Types of accounts include and are defined as:

1. **Personal**

- a. People who represent their own views and interests and use their account for fun, to read the news, and socially interact (Uddin, Imran, and Sajjad 2014, 2)

2. **Professional**

- a. Share information about their area of expertise and may include businesses who use the platform to market their products and reach new audiences (Uddin, Imran, and Sajjad 2014, 2)

3. **Bots**

- a. Automated “... online agents that are used to intervene in political discourse online” and are intentionally created to circulate disinformation, malicious content, and flood the notifications of users (Dubois and McKelvey 2019, 28), (Marlow, Miller, and Roberts 2021).

4. **Spammers**

- a. Intentionally “spread malicious content” (Singh, Bansal, and Sofat 2018, 2).
- b. May show up as one individual who has created multiple accounts to bolster their ability to send harassing content (Mantilla, 2015).

5. Anonymous

- a. An individually intentionally hiding information about their identity and location for the purposes of using the platform maliciously. May include not adding a profile picture or using an obscured, random picture and not revealing their actual name, age, gender, and other identifying characteristics (Trifiro et al. 2021).
- b. Hide behind fake usernames and “... express sexist, racist, or homophobic views with little fear of being held accountable for their comments” (Wagner 2020, 8).

6. Suspended/Deleted

- a. Accounts found to be in violation of Twitter’s Terms of Services and temporarily or permanently removed from the platform.

3 Other Guidelines

3.1 Retweets

Retweets are not to be coded, only responses to Tweets that are initiated from the accounts of one of the three female politicians. A simple retweet or quote Tweet (retweet with comment) would not count, as the Tweet being shared was not originally authored by one of female politicians.

3.2 Threaded Responses

Threaded responses occur when one user posts a connected series of Tweets which contribute towards building the same message or narrative. They may be numbered as 1, 2, 3, ... out of the total number of responses.

For the purposes of this study, each Tweet in a threaded response will be analyzed and coded as a separate Tweet.

3.3 Media

Twitter defines any gifs, memes, images, and videos as media (Developer 2023). Media embedded in Tweets should be considered as part of the written text of the Tweet and coded accordingly, given that hateful speech can also be expressed through images and symbols (Carlson 2021). With increasing doctoring and creation of media for the purpose of deception and political gain, any media embedded in Tweets should be analyzed for manipulation and machine-learning generated deepfakes, which are compelling fake pictures and videos (Farid 2019), (Appel and Prietzel 2022).

For Tweets that contain a picture without additional text, the nature, context, and suggestiveness of the picture should be analysed to determine the severity of harassment (Quan-Haase and Sloan 2017). Emojis, if included in a Tweet, should be interpreted as both a picture and text, analysing the rest of the Tweet for context (Swartz and Crooks 2020). Moreover, hashtags will be coded in the same manner as regular text in Tweets (Quan-Haase and Sloan 2017).

3.4 Hate Speech

It is anticipated that the quantity of hate speech detected in this sample of Tweets will be quite low, given that these Tweets are from 2019 when Canadian politics was slightly less polarized on Twitter and Twitter employed their content moderation rules more frequently. However, given changes to Canadian politics and the affordances of Twitter, this is likely to change in the full-scale version of this study.

References

- Appel, Markus, and Fabian Prietzel. 2022. "The Detection of Political Deepfakes." *Journal of Computer-Mediated Communication* 27 (4): zmac008.
- Carlson, Caitlin Ring. 2021. *Hate Speech*. MIT Press.
- Chang, Winston, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. 2023. *shiny: Web Application Framework for R*. <https://shiny.posit.co/>.
- Chetty, Naganna, and Sreejith Alathur. 2018. "Hate Speech Review in the Context of Online Social Networks." *Aggression and Violent Behavior* 40: 108–18.
- Developer, Twitter. 2023. "Follow, Search, and Get Users." Twitter Developer. <https://developer.twitter.com/en/docs/twitter-api/v1/accounts-and-users/follow-search-get-users/overview>.
- Dubois, Elizabeth, and Fenwick McKelvey. 2019. "Political Bots: Disrupting Canada's Democracy." *Canadian Journal of Communication* 44 (2): PP–27.
- Farid, Hany. 2019. *Fake Photos*. MIT Press.
- Firke, Sam. 2023. *janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- Jansen, Bernard J, Mimi Zhang, Kate Sobel, and Abdur Chowdury. 2009. "Twitter Power: Tweets as Electronic Word of Mouth." *Journal of the American Society for Information Science and Technology* 60 (11): 2169–88. <https://doi-org.myaccess.library.utoronto.ca/10.1002/asi.21149>.
- Krippendorff, Klaus. 2011. "Agreement and Information in the Reliability of Coding." *Communication Methods and Measures* 5 (2): 93–112.
- Mantilla, Karla. 2013. "Gendertrolling: Misogyny Adapts to New Media." *Feminist Media Studies* 39 (2): 563–70.

- Marlow, Thomas, Sean Miller, and J Timmons Roberts. 2021. "Bots and Online Climate Discourses: Twitter Discourse on President Trump's Announcement of US Withdrawal from the Paris Agreement." *Climate Policy* 21 (6): 765–77.
- Nadim, Marjan, and Audun Fladmoe. 2021. "Silencing Women? Gender and Online Harassment." *Social Science Computer Review* 39 (2): 245–58.
- Quan-Haase, Anabel, and Luke Sloan. 2017. "The SAGE Handbook of Social Media Research Methods." *The SAGE Handbook of Social Media Research Methods*, 1–728.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Saleem, Haji Mohammad, Kelly P Dillon, Susan Benesch, and Derek Ruths. 2017. "A Web of Hate: Tackling Hateful Speech in Online Social Spaces." *arXiv Preprint arXiv:1709.10159*.
- Singh, Monika, Divya Bansal, and Sanjeev Sofat. 2018. "Who Is Who on Twitter—Spammer, Fake or Compromised Account? A Tool to Reveal True Identity in Real-Time." *Cybernetics and Systems* 49 (1): 1–25.
- Swartz, Melanie, and Andrew Crooks. 2020. *Comparison of Emoji Use in Names, Profiles, and Tweets*. IEEE. <https://www.computer.org/csdl/proceedings-article/icsc/2020/633200a375/1iffz19otJS>.
- Theocharis, Yannis, Pablo Barberá, Zoltán Fazekas, and Sebastian Adrian Popa. 2020. "The Dynamics of Political Incivility on Twitter." *Sage Open* 10 (2). <https://doi.org/10.1177/2158244020919447>.
- Trifiro, Briana M, Sejin Paik, Zhixin Fang, and Li Zhang. 2021. "Politics and Politeness: Analysis of Incivility on Twitter During the 2020 Democratic Presidential Primary." *Social Media+ Society* 7 (3): 20563051211036939.
- Uddin, Muhammad Moeen, Muhammad Imran, and Hassan Sajjad. 2014. "Understanding Types of Users on Twitter." *arXiv Preprint arXiv:1406.1335*.
- Vickery, Jacqueline Ryan, and Tracy Everbach. 2018. *Mediating Misogyny: gender, Technology and Harassment*. Springer.
- Wagner, Angelia. 2020. "Tolerating the Trolls? Gendered Perceptions of Online Harassment of Politicians in Canada." *Feminist Media Studies* 22 (1): 32–47.
- Zavattaro, Staci M, P Edward French, and Somya D Mohanty. 2015. "A Sentiment Analysis of US Local Government Tweets: The Connection Between Tone and Citizen Involvement." *Government Information Quarterly* 32 (3): 333–41. <https://doi-org.myaccess.library.utoronto.ca/10.1016/j.giq.2015.03.003>.