THINK BETTER, YOU DUMBASS: ONLINE HATEFUL SPEECH AS EPISTEMIC VIOLENCE

Esteban Morales
University of Groningen

Jaigris Hodson
Royal Roads University

Victoria O'Meara
University of Leicester

Introduction

Online violence and abuse are common for many users —and particularly so in spaces where people seek to engage in dialogue, such as in comment sections on online news sites (Zannettou et al., 2020). Previous scholarship has noted that, far from being innocuous, online violence in comment sections harms the people targeted (Harlow, 2015), undermines the credibility of the news site (Masullo et al., 2023), and results in an overall erosion of public discourse (Masullo Chen & Lu, 2017). These consequences are exacerbated by social structures of power that are sustained through these comments, which tend to favour patriarchal, racial, and nationalist ideologies (Harlow, 2015; Salgado et al., 2024; Wolfgang, 2022).

In this paper, we focus on one particular impact of toxic speech on digital platforms: its epistemological implications. Here, the work of Galpin and Vernon (2024) has showcased how violence on digital platforms can inhibit people's ability to participate freely in the public sphere. Expanding on their work, we explore how toxic speech functions as a manifestation of epistemological violence—that is, as efforts to erase particular ways of thinking (Colombo, 2020). To achieve this, we first explore the mechanisms through which epistemic violence is enacted, and then, recognizing that these epistemic enclosures are not felt equally by everybody, we center how identity comes at the forefront of the uses of toxic speech to enact epistemic violence.

Methods

We focus on an analysis of online commentary on The Conversation Canada—a news organization that publishes articles written by academic researchers, translating scholarly insights for a general audience (The Conversation Canada, n.d.). As part of an

ongoing collaboration between the authors and the news organization, The Conversation Canada donated a comment database containing all the public user interactions in the stories published between 2020 and the date of data collection—March 2024. This database included a total of 31,682 comments on 3,592 articles.

To identify the messages that displayed hateful and toxic speech, we relied on toxicity analysis via the Perspective API, which we accessed through Communalityc (Gruzd & Mai, 2023). The Perspective API uses machine learning to mark online comments from 0 (non-toxic) to 10 (very toxic)—defining toxicity as "a rude, disrespectful, or unreasonable comment that is likely to make you leave a discussion" (Perspective API, n.d.). Through this process, we identified 390 toxic comments in 199 articles.

Recognizing that there are limitations to automated toxicity analysis (e.g., Oh & Downey, 2024), we approach this analytical stage as a filtering tool to answer our research question. In this sense, we conducted a manual thematic analysis of the toxic comments (Braun et al., 2019). We qualitatively examined the role of hateful speech in enacting epistemic violence, closely exploring the topics where violence was likely to happen and in what way identity was positioned in the data.

Findings

We identified four mechanisms: insulting, labelling, ridiculing, and dehumanizing. The purpose of all these four mechanisms is to erase 'the wrong way' of thinking about the topic at hand and enforce an epistemic closure.

Insulting. The first mechanism refers to how toxic speech is directed at the people voicing the perspective as a way to shut down these perspectives. For instance, insults such as "you are one dumb ignorant BITCH," and "Get over it you fuckwits!" aim to discourage people from expressing or aligning with particular epistemic positioning.

Labelling. The second mechanism concerns the use of toxic speech to categorize epistemic stances within larger societal constructs—and then dismiss them. For example, a user commenting on a story about donation bins in Canadian supermarkets called the author "a left-wing Neo liberal imbecile." Such comments seek to align the speaker and their perspective with certain political and ideological identity groups in an effort to discredit them.

Ridiculing. The third mechanism seeks to undermine epistemic stances by mocking them—thus making the perspective not serious. These efforts of ridiculing can be seen at two levels: (1) They can be directed at the epistemological perspective itself (e.g., "That is the most stubid [sic], idea[...]" or (2) at the person putting forward the stance (e.g., "The author must be on drugs"). These forms of ridiculing seek to undermine the confidence of people aligning with a specific epistemological stance and exclude them from meaningfully engaging in public conversations (McSwiney & Sengul, 2024).

Dehumanizing. The final mechanism concerns dehumanization—that is, depriving those holding particular perspectives of human qualities. This mechanism is visible in comments that characterize the speaker as an animal, as in "My family arrived in

Canada long before a pig like you did [...]" as well as in comments that compare the speaker to refuse, as in "[...] you're a lawyer, so I guess we should just expect you to act like the subhuman, piece of trash that you are." By removing the human qualities of those holding particular ways of thinking, these uses of toxic speech erase the possibility of dialogue (Tutkal, 2023).

While these four mechanisms were seen across the comments, we found numerous instances where identity was the focus of toxic comments. Indeed, we saw evidence of this across race (e.g., "That is such bullshit. Black women are hypochondriacs"), gender (e.g., "It's written by a woman and gender studies professor, of course it's rubbish"), partisan politics (e.g., "The democrats are morally depraved animals") and ethnicity (e.g., We need more American made weapons so we can mass murder arabs and latinos"). These findings showcase how online abuse is operationalized within larger structures of power as mechanisms for establishing moral boundaries concerning who is allowed (and not allowed) to inhabit and speak in the public sphere (Galpin & Vernon, 2024).

Conclusion
Our work illustrates four different mechanisms (insulting, labelling, ridiculing, and dehumanizing) through which hateful language incites epistemic closure. Through these four mechanisms, specific ways of thinking are attacked and discouraged. In line with previous work (Galpin & Vernon, 2024; Salgado et al., 2024; Wolfgang, 2022), our work showcases how hateful speech is used to reinforce efforts of mono-universalization of truth across ideas that tend to follow partisan, racial, patriarchal, and nationalist ways of thinking. This paper contributes to a better understanding of the negative impacts of online toxicity (e.g., Coombs, 2021; Jane, 2018; Recuero, 2024), showing the intent of online toxicity is often to bully others (particularly diverse others) into conformity. While this paper has focused on toxic comments from a news site, future work could explore other public forums (such as social media platforms). Similarly, given that our analysis focuses on the words of perpetrators (rather than the experiences of targets), future studies could further explore the epistemological consequences of online hateful speech.

Foregrounding the epistemological consequences of hateful speech calls for new approaches to address pushes for a mono-universalization of truth. As long noted by post-colonial scholars (Spivak, 2023), the solution to epistemic violence is not merely to stamp out violent behaviours. Instead, the goal is to ensure pluralism of worldviews, whereby various ways of thinking find a place to exist and dialogue (Hoggan-Kloubert & Hoggan, 2023; Novis-Deutsch, 2020). Following this line of thought and inspired by the work of Arturo Escobar (2018), an urgent task is to imagine and design an epistemological ground of digital conversations where many epistemological stances fit.

References
Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic analysis. In P. Liamputtong (Ed.), Handbook of Research Methods in Health Social Sciences (pp. 843–860). Springer Singapore. https://doi.org/10.1007/978-981-10-5251-4_103

Colombo, M. (2020). Who is the "other"? Epistemic violence and discursive practices. Theory & Psychology, 30(3), 399–404. https://doi.org/10.1177/0959354320923758

Coombs, E. (2021). Human rights, privacy rights, and technology-facilitated violence. In J. Bailey, A. Flynn, & N. Henry (Eds.), The Emerald International Handbook of Technology-Facilitated Violence and Abuse (pp. 475–491). Emerald Publishing Limited. https://doi.org/10.1108/978-1-83982-848-520211036

Escobar, A. (2018). Designs for the Pluriverse: Radical Interdependence, Autonomy, and the Making of Worlds. Duke University Press.

Galpin, C., & Vernon, P. (2024). Post-truth politics as discursive violence: Online abuse, the public sphere and the figure of 'the expert.' The British Journal of Politics and International Relations, 26(2), 423–443. https://doi.org/10.1177/13691481231202641

Gruzd, A., & Mai, P. (2023). Communalytic – A Research Tool For Studying Online Communities and Online Discourse. [Computer software]. https://communalytic.org/

Harlow, S. (2015). Story-Chatterers Stirring up Hate: Racist Discourse in Reader Comments on U.S. Newspaper Websites. Howard Journal of Communications, 26(1), 21–42. https://doi.org/10.1080/10646175.2014.984795

Hoggan-Kloubert, T., & Hoggan, C. (2023). Post-Truth as an Epistemic Crisis: The Need for Rationality, Autonomy, and Pluralism. Adult Education Quarterly, 73(1), 3–20. https://doi.org/10.1177/07417136221080424

Jane, E. A. (2018). Gendered cyberhate as workplace harassment and economic vandalism. Feminist Media Studies, 18(4), 575–591. https://doi.org/10.1080/14680777.2018.1447344

Masullo Chen, G., & Lu, S. (2017). Online Political Discourse: Exploring Differences in Effects of Civil and Uncivil Disagreement in News Website Comments. Journal of Broadcasting & Electronic Media, 61(1), 108–125. https://doi.org/10.1080/08838151.2016.1273922

Masullo, G. M., Tenenboim, O., & Lu, S. (2023). "Toxic atmosphere effect": Uncivil online comments cue negative audience perceptions of news outlet credibility. Journalism, 24(1), 101–119. https://doi.org/10.1177/14648849211064001

McSwiney, J., & Sengul, K. (2024). Humor, Ridicule, and the Far Right: Mainstreaming Exclusion Through Online Animation. Television & New Media, 25(4), 315–333. https://doi.org/10.1177/15274764231213816

Novis-Deutsch, N. (2020). Pluralism as an antidote to epistemic violence in psychological research. Theory & Psychology, 30(3), 408–413. https://doi.org/10.1177/0959354320928116

Oh, D., & Downey, J. (2024). Does algorithmic content moderation promote democratic discourse? Radical democratic critique of toxic language AI. Information, Communication & Society, 1–20. https://doi.org/10.1080/1369118X.2024.2346531

Perspective API. (n.d.). Home. Retrieved February 19, 2025, from
https://perspectiveapi.com/

Recuero, R. (2024). The platformization of violence: Toward a concept of discursive
toxicity on social media. Social Media + Society, 10(1), 20563051231224264.
https://doi.org/10.1177/20563051231224264

Salgado, S., Gil Zúñiga, H., Da Silva, P. A., Biscaia, A., Coimbra, M. E., Martins, B., &
Francisco, A. (2024). Assessing the Prevalence and Predictors of Incivility in
Online News Comments Across Six Countries. Journalism Practice, 18(2),
224–241. https://doi.org/10.1080/17512786.2023.2246431

Spivak, G. C. (2023). Can the Subaltern Speak? Imperialism, 171–219.

The Conversation Canada. (n.d.). Who we are. The Conversation. Retrieved February
19, 2025, from https://theconversation.com/ca/who-we-are

Tutkal, S. (2023). The role of dehumanization in legitimation and delegitimation of state
violence in Colombia. Culture & Psychology, 1354067X2311647.
https://doi.org/10.1177/1354067X231164795

Wolfgang, J. D. (2022). When Fringe Hate Goes Mainstream: How White Nationalist
Discourse Manifests in Online News Commenting. Journal of Communication
Inquiry, 46(2), 117–137. https://doi.org/10.1177/01968599211040840

Zannettou, S., Elsherief, M., Belding, E., Nilizadeh, S., & Stringhini, G. (2020).
Measuring and Characterizing Hate Speech on News Websites. Proceedings of
the 12th ACM Conference on Web Science, 125–134.
https://doi.org/10.1145/3394231.3397902