



Selected Papers of #AoIR2018:
The 19th Annual Conference of the
Association of Internet Researchers
Montréal, Canada / 10-13 October 2018

HATRED OF/AND DEMOCRACY: THE POLITICAL CONTRADICTIONS OF REDDIT'S MODERATION STRUCTURE

Dr. Trevor Smith
Carleton University

In February 2018, chat transcripts of a conversation between the moderators of Reddit subforum r/canada were leaked by one of the participants. These transcripts show a debate about whether or not to ban a user with a notoriously racist posting history. In the course of this conversation, one of the moderators reveals he is "slowly becoming a white nationalist." This moderator scandal capped a long decline in the quality of comments posted to r/Canada which has been repeatedly remarked upon by users.

The fall of r/canada into becoming a safe haven for hate presents an interesting case through which to examine the fundamental contradictions in the way Reddit's communities are governed. Despite being one of the few nascent sites of possible political space, Reddit retains a thoroughly anti-political governance structure. If we think of Reddit's subreddits in terms of governmental structures, they tend to resemble authoritarian dictatorships more than the user-driven democracy that Reddit likes to present itself as to the public. Part of this disconnect is driven by the contradiction between the role of human moderation and Reddit's user-driven public relevance algorithm.

Reddit once billed itself in explicitly democratic terms with its slogan of "today's headlines, chosen by you, not editors." Reddit's upvote/downvote system provides a mechanism for users to help shape what is more or less visible on the site. Yet Reddit does not seem to trust their own algorithm to do its job, as each subreddit is governed by moderators who are vested with complete authority over the content of a subreddit. Like in a hereditary monarchy, their authority is based on their ability to have either come first or to have been appointed by someone who came before. Furthermore, the arguments provided to legitimize this form of political authority tend to be a rehash of Plato's arguments against democracy.

Here is where interesting contradictions arise, as if the algorithm did its job, hate should fall to the bottom and ultimately be buried by downvotes. But Reddit does not trust the people to vote correctly, and thus the human layer of moderation is given complete

Suggested Citation (APA): Smith, T. (2018, October 10-13) *Hatred Of/And Democracy: The Political Contradictions of Reddit's Moderation Structure*. Paper presented at AoIR 2018: The 19th Annual Conference of the Association of Internet Researchers. Montréal, Canada: AoIR. Retrieved from <http://spir.aoir.org>.

authority over content. However, placing large communities in the hands of a few individuals with arbitrary authority has been a breeding ground for abuse.¹ Typical analysis² positions online hate as stemming from a lack of accountability and consequences, and tends to result in the demand for more control of online spaces. In the case of r/canada, the moderators who were not so sympathetic to hate seemed to have misdiagnosed the problem of increased racism and hate in two ways. The first was the idea that the subreddit was suffering from a lack of moderation, leading to the addition of more moderators (including, strangely, moderators of the rival r/metacanada subreddit which is dedicated to mocking the users of r/Canada). The second misdiagnosis came in the form of believing the problem of hate stemmed from a lack of politeness rather than a lack of civility.³ Many users reported having comments removed in which they impolitely challenged racism, while the uncivil racist comments were allowed to stay, sewing confusion among the user base over moderation policies.

The argument that hate stems from a lack of control and a lack of consequences betray what Jacques Rancière⁴ described as the very common hatred of democracy. Despite the fact that everyone is willing to pledge allegiance to the word democracy, in reality the practice of democracy is met with general disdain. Drawing on Rancière's analysis of the perception of democracy, I argue that the rise of hate on r/Canada is symptomatic of the lack of trust in the people that is characteristic of the structure of many online communities which are dedicated to political discussion.

While Reddit could be positioned as having a dual structure with a democratic algorithm and authoritarian moderators, the actual interplay between the two elements of moderation tend to work together to undermine the possibility of political discourse. With moderators prone to personal bias, extreme reactions, and lacking accountability to their communities, algorithmic moderation is usually posited as the alternative. In the case of Reddit, their "democratic" algorithm is actually fundamentally anti-political.

As theorists of agonistic democracy⁵ argue, politics is inherently conflictual and exists because of disagreement. Reddit's public relevance algorithm works well on subreddits dedicated to fluffy animals where no one wants to see mean comments, but breaks when the discussion turns political. Up and down votes become proxies for agree/disagree, and comments that are politically valuable since they generate discussion can end up hidden because the algorithm punishes unpopularity and controversy. Reddit's algorithm produces an echo chamber effect, or in Reddit parlance a "circlejerk," by hiding legitimate discussion in favour of empty yet popular platitudes.

¹ Alfonso, Fernando. "Meet the Reddit Power User Who Helped Bring down R/technology." The Daily Dot, April 22, 2014. <http://www.dailydot.com/politics/reddit-maxwellhill-moderator-technology-flaw/>.

Auerbach, David. "Does Reddit Have a Transparency Problem?" Slate, October 9, 2014. http://www.slate.com/articles/technology/technology/2014/10/reddit_scandals_does_the_site_have_a_transparency_problem.html.

² Suler, John. "The Online Disinhibition Effect." *CyberPsychology & Behavior* 7, no. 3 (2004): 321–26. Dreyfus, Hubert L. *On the Internet*. London: Routledge, 2008.

³ Papacharissi, Zizi. "Democracy Online: Civility, Politeness, and the Democratic Potential of Online Political Discussion Groups." *New Media & Society* 6, no. 2 (2004): 259–83.

⁴ Rancière, Jacques. *Hatred of Democracy*. Translated by Steve Corcoran. London: Verso, 2009.

⁵ Mouffe, Chantal. *On The Political*. London: Routledge, 2005.

Honig, Bonnie. *Political Theory and the Displacement of Politics*. Ithaca: Cornell University Press, 1993.

In this sense Reddit suffers from a doubly anti-political structure. Moderators act like feudal lords defending their territory and the algorithmic structure of the site suppresses political disagreement. Here again, Rancière's claims about a hatred of democracy lingering under the surface of claims of fidelity to democracy are instructive. Thinking back to r/Canada, the double problem of anti-politics shows itself, as moderators sympathetic to hate remove users who challenge the hate, while moderators who would prefer not have hate can only think in terms of top-down solutions. The idea of fostering a democratic culture in which hate is met with persistent rebuke is not even on the radar of possibilities.

If human moderators and algorithmic moderation are both "worse," how can this debate about how to foster genuine political discussion online be solved? My argument is to turn back to politics itself. The problem is not at the level of overly biased humans vs. overly objective algorithms but really a question of political theory. When users of a political discussion website, such as the many subreddits dedicated to the topic, feel like they are merely peasants invading the realm of a lord with arbitrary powers backed by a system of law based on contradictions, it's no wonder that the peasants are less than civil.

I suggest that online politics requires a shift toward democratic structures that allow the users to be direct participants in how the community is governed. By investing users with a say in the structure of a site, a culture of ownership and buy-in can be fostered in which users feel like they are not merely guests or intruders in someone else's space, but are responsible themselves for the quality of the space. Suggestions such as this are often met with skepticism due to the claim that such a system would be unable to deal with trolls and hate. Yet the arguments against online democracy once again mirror Plato's critique of democracy as both relying on universal impersonal laws which lack the care and concern of a human moderator who can guide the flock and as a situation governed by individuals motivated only by their own pleasure. Ironically, as Rancière points out, the critiques of online democracy are the same as the critiques of ancient democracy, which we are told repeatedly is an unworkable model in our modern technological context.

Returning to the issue of hate on r/Canada, perhaps the original evil is not the unruly and conflicted masses as is so often is blamed, but instead the figure of the moderator rallying the mob around a singular authority promoting hate. I flesh this position out with further references to the history of the decline of r/Canada, in which a sudden influx of moderator activity left many users baffled about the rules and set the stage for hate to become prevalent, demonstrating the point that hate is a function of top down control and not of its absence.