DISPLACED DISCUSSION: THE IMPLICATIONS OF REDDIT QUARANTINE AND THE MOVEMENT OF THEREDPILL TO SELF-HOSTING

Luc S. Cousineau
University of Waterloo

We rely on social media companies to manage content on our behalf. From hosting our pictures to curating our feeds, these companies do as much (or more) to manage content we do not see as the content that we do see. For years, writers and researchers like Sarah Roberts (2019), Safiya Noble (Noble, 2018), and Tarlton Gillespie (2018) have written about commercial content moderation, biased algorithms, and the embedded politics in platforms. They have exposed how companies manage us and our content, often unbeknownst to users. But what happens to communities managed by site moderation schemes when they remain committed to operating as a group? Especially when those groups hold views steeped in misogyny and male supremacist discourse?

Men's groups online have long been users of digital technologies to help them unite geographically disparate membership (Borton, 2017; Munn, 2019). Using early text-based communication services and forums, men's rights groups grew their numbers through community and direct connections. However, maintaining websites and forums requires human, economic, and commitment capital, so it is far easier (and cheaper) to use existing hosted platforms, but gives up some measure of control (Plantin et al., 2016).

Reddit.com is a social media site with huge volumes of content hosted on their platform, and imported from other providers. In moderating that content, Reddit must consider its users, investors, corporate oversight, and its own brand identification as a place for free speech. At the user level, Reddit engages community moderators; uncompensated volunteers that manage communities and rules, delete posts, and ban users as a way to maintain individual sub-community standards. At the site-wide level, Reddit employs two

types of broad content management over its sub-communities: banning and quarantine. Reddit has banned communities (called subreddits) from the site (Reddit does not publish a list, but some of note include: r/the_donald, r/fatpeoplehate, r/watchpeopledie, and r/beatingwomen). However, Reddit will often quarantine bad-actor communities before instituting a ban. Quarantine leaves these communities operational on the site, but hides them from regular users with an extra click-through acknowledgement, and stops their content from appearing in searches. This approach is Reddit's way of censuring the group without censoring the group. Quarantined groups can still post, act, and talk, but tucked away in a sort of bad-actor closet.

Reddit is an established site of counterpublic community action. Sub-Reddits like r/menrights, r/DegradeEbonyThots, and (the now defunct) r/the_donald expound, in turn, anti-feminist, deeply racist and misogynist, and (sometimes) violent right-wing ideologies. Along with the various posts, links, and outside content, Reddit communities support millions of words of user discussion, including hate and violence. While bans have been shown to reduce hate speech (Chandrasekharan et al., 2017), quarantine and individual user sanctions seem to have mixed effects on user behaviour (Myers West, 2018; Shen & Rose, 2019). However, sub-communities are policed, at least in part, by their exposure to the community at large (Hui & Buchegger, 2009). So what happens to communities quarantined or otherwise limited on Reddit? Some may turn toward other platforms, including self-hosting.

This work uses data from a digital ethnography of r/TheRedPill, a subreddit dedicated to "discussion of sexual strategy in a culture increasingly lacking a positive identity for men." Data collection included community post and comment observation, and using website capture tools to systematically capture the landing page and top-rated posts over a period of three months. Analysis used a technofeminist lensing (Wajcman, 2004) and community theory (Cohen, 2001), and included review of the captured pages and posts/comments, as well as researcher field notes from community observation.

Data collection began in fall 2019, and continued to summer 2020. r/TheRedPill was quarantined in September of 2018, and has since engaged in regular discussion about the movement of community discussion and forums away from Reddit to their own site, www.trp.red. The push for this move was renewed in October 2019 with the introduction of new community standards by Reddit, and again in February 2020 with minor changes to the way individuals can (and can not) access content on the subreddit.

The development of www.trp.red as a forum and discussion site to mirror Reddit's functionality is linked, at least in part, to the limiting of the community through Reddit's quarantine scheme. Given the dispersed nature of the internet, coupled with the quasi-anonymous nature of Reddit, it is difficult to prove that users that choose to leave Reddit migrate elsewhere, however, in the case of r/TheRedPill, users and moderators actively post about and encourage the move to their own platform, and share usernames

between the sites. While a r/TheRedPill exodus might meet the needs of Reddit as a business (having secured hundreds of millions of dollars, and is investigating an IPO), progressively limiting communities and pushing them out also limits the potential dampening of contentious speech and regressive social attitudes that comes with wider public pressure to conform. When we push contentious communities out of the public eye, there are fewer checks and balances to limit moves towards extremism.

Beyond a discussion of simply allowing versus limiting speech, the case of r/TheRedPill provides an opportunity to engage in a discussion about whether progressive sanction is the right way to manage the intersection of counterpublic views with the needs of tech firms, especially in a sector where the division between public and private feels in flux. This paper hopes to further that discussion using a community that is less generally associated with hate speech, and can therefore exist on the margins of acceptability.

References

Borton, D. (2017). *Masculinity and Misogyny in the Online Locker Room: Subcultural Identity Work in the Title: Men's Rights Movement* [M.A., Indiana University of Pennsylvania]. http://search.proquest.com/docview/1944007399/abstract/CA6DF543874A4821PQ/1

Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J., & Gilbert, E. (2017). You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 1–22. https://doi.org/10.1145/3134666

Cohen, A. P. (2001). *The symbolic construction of community* (Transferred to digit. pr). Routledge.

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Hui, P., & Buchegger, S. (2009). Groupthink and Peer Pressure: Social Influence in Online Social Network Groups. *2009 International Conference on Advances in Social Network Analysis and Mining*, 53–59. https://doi.org/10.1109/ASONAM.2009.17

Munn, L. (2019). Alt-right pipeline: Individual journeys to extremism online. *First Monday*, *24*(6). https://doi.org/10.5210/fm.v24i6.10108

Myers West, S. (2018). Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, *20*(11), 4366–4383. https://doi.org/10.1177/1461444818773059

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2016). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, *Journal Article*, 1461444816661553.

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press.

Shen, Q., & Rose, C. (2019). The Discourse of Online Content Moderation: Investigating Polarized User Responses to Changes in Reddit's Quarantine Policy. *Proceedings of the Third Workshop on Abusive Language Online*, 58–69. https://doi.org/10.18653/v1/W19-3507

Wajcman, J. (2004). *TechnoFeminism*. Polity.