The Controversy over Social Media Recommendation Algorithms

A Sociotechnical Research Paper
presented to the faculty of the
School of Engineering and Applied Science
University of Virginia

by

Conor Monaghan

May 12, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Conor Monaghan

Sociotechnical advisor: Peter Norton, Department of Engineering and Society

Introduction

From post-it notes that assist with memory to the devices that clean our clothes and wash our dishes, humans constantly invent methods to offload tasks and responsibilities. Now humans have invented something that substitutes the need to search for and collect information—the recommendation algorithm. These algorithms form the backbone of social media sites such as Twitter, Facebook, and YouTube. And since 69 percent of adults in the U.S. use Facebook, and 73 percent use YouTube, these algorithms have power to shape public thought (Perrin, Anderson 2020). That power has led to politicians concerned with issues such as free speech, bias, and national security, while former tech insiders worry about extremism and the algorithm's addictive nature. Groups like these want to add new regulations to recommendation algorithms and content moderation, while companies such as Facebook and Twitter want to continue their use as is. How do companies, politicians, and advocacy groups compete to change or retain regulations around recommendation algorithms and content moderation on social media platforms? Social media corporations appeal to free speech, self-regulation, and user satisfaction when defending their algorithms, while critics raise concerns over bias and extremism.

Review of Research

Johnson et al. (2019) used mathematical analysis to show that "extreme sub-populations will likely be enhanced over time by new social media algorithms designed to reduce division." Bakshy et al. (2015) found that on Facebook, the more ideological content is, the likelier it is to be shared, and "the News Feed ranking algorithm sorts these articles," which influences what individuals read. Covington et al.'s (2016) description of YouTube's recommendation algorithm

shows that the priority is to increase watch time and other retention data. This increases user time on the site and can amplify extreme content that engages users.

Lacy-Nichols et al. (2019) contend that the Australian Beverage Council avoided government regulation using a combination of funded research and self-regulation to decrease the "perceived urgency of the problem." In a study of self-regulation in food marketing to children in the US, Fleming-Milici and Harris (2020) noted that "a major barrier to effective self-regulation is that food and beverage companies' primary goal – to generate profits – is antithetical to improving the health of children," especially when "a company's major competitors do not also reduce their marketing of unhealthy products." This suggests social media companies, which profit off of the time users spend on their site, would have little incentive to effectively self-regulate if it reduced user activity. There is also large incentive to not self-regulate when others are, since users that spend less time on the less addictive sites would spend that time on the more addictive ones.

Cumming et al. (2005) found that cigarette companies defended their addictive cigarette design by arguing the cigarettes were designed for the customer's satisfaction, not nicotine levels, and that the companies "have worked with the public health community to make the safest cigarette acceptable to the smoker." The companies also contended that for a cigarette to be safe, "it must be acceptable to a large number of consumers," or else consumers would not switch to the safer cigarettes.

Social Media Companies and Regulation

Social media companies often argue that what is best for the public is also best for their platforms. Google's CEO, Sundar Pichai, maintained Google had no political bias (*Big Tech*

Company's Liability Shield, 2020), as "to do otherwise would be contrary to both our business interests and our mission, which compels us to make information accessible to every type of person." In a congressional hearing titled *Twitter: Transparency and Accountability*, Twitter's CEO, Jack Dorsey (2018), likened Twitter to a "digital public square," and said that "a key driver of a thriving public square is the fundamental human right of freedom of opinion and expression." Due to this, "Twitter cannot rightly serve as public square if it's constructed around the personal opinions of its makers" and "will always default to open and free exchange." Facebook CEO Mark Zuckerberg (*Amazon/Facebook/Google/Apple Dominance*, 2020) said Facebook's mission is "connecting people around the world," and the company supports this by selling ads. By framing their companies not as profit-driven, but as ones that operate to fulfill a purpose, social media corporations align their perceived interests with public interests, and portray monetary decisions as necessary to achieve their goals.

One of the ways social media companies downplay the need for regulation is by highlighting their ongoing efforts to improve their services and algorithms. Addressing bias, Jack Dorsey (*Twitter: Transparency and Accountability*, 2018) said that Twitter was "collaborating with the non-profit research center Cortico and the Massachusetts Institute of Technology Media Lab on exploring how to measure aspects of the health of the public sphere." Zuckerberg (*Amazon/Facebook/Google/Apple Dominance*, 2020) highlighted "35,000 people working on safety and security," and said Facebook was "funding new technologies to tackle emerging threats like deepfakes." If politicians are assured that social media companies are working to improve themselves, they may be less likely to pass legislation, even if those companies have little improvement.

Social media companies suggest that the best way to fix problems with their site is to create better algorithms. Twitter plans to reduce abuse by creating "technology to recognize [abuse] before people have to do the reporting themselves" (*Twitter: Transparency and Accountability*, 2018). Facebook's plan to reduce extremism on the platform is to use systems that "are able to detect when there's potential issues," and is working "with law enforcement and the intelligence community" to be able to detect potential threats (*Big Tech Company's Liability Shield*, 2020). During the pandemic, YouTube relied "more on technology to quickly identify and remove content that violates our Community Guidelines so that our teams that review content could safely remain at home" (YouTube Team, 2020). Companies that prescribe improved algorithms as the solution to problems can give politicians the impression that issues with social media are best addressed within their company.

Social media companies can attribute issues on their site to defective algorithms or polices that can be easily fixed. In the hearing *Twitter: Transparency and Accountability*, Jack Dorsey (2018) admitted Twitter was unfairly filtering accounts from search suggestions, and said "Our technology was using a decision-making criteria that considers the behavior of people following these accounts. We decided that wasn't fair and we corrected it." When Zuckerberg was asked what Facebook will do to make sure militia groups are removed from the site, he said Facebook had "strengthened our policies" to produce "fewer mistakes" (*Big Tech Company's Liability Shield*, 2020). In the same hearing, when Dorsey was asked why Twitter removed a story about Hunter Biden, he said: "we updated our policy and our enforcement within 24 hours." By portraying issues as defects in an algorithm or policy that can be quickly fixed, companies can assure politicians that bias can be easily quantified and eliminated.

Social media companies stress that their systems are made to provide the user with the best experience. In 2018, Facebook changed its ranking algorithm to prioritize posts from friends and family over other media. The company stressed that the updates were meant to connect "people to meaningful posts" and "spark conversations and meaningful interactions between people" (Mosseri, 2018). Dorsey's testimony (*Twitter: Transparency and Accountability*, 2018) emphasized that Twitter's algorithms are made to provide relevant content for each user, and individuals are shown posts in ranked order because "this feature creates a better experience for people using Twitter by showing people the Tweets they might find most interesting first." YouTube (Meyerson, 2012) started focusing on recommendations that "increase the amount of time that the viewer will spend watching videos on YouTube, not only on the next view, but also successive views thereafter." They justified this by saying the changes mean viewers are "happier with the content they've found," and that now "creators are attracting more engaged audiences." Companies can promote changes that increase user activity on their site as beneficial to users.

Social media companies promote transparency as a way to gain public trust. In his testimony, Zuckerberg touted the numbers of posts Facebook had removed for illegal content (*Big Tech Company's Liability Shield*, 2020), saying "We report these numbers as part of our Transparency Reports, and we believe all other major platforms should do the same." Congress should work on "making content moderation systems more transparent." Dorsey, in his testimony (*Twitter: Transparency and Accountability*, 2018), said "We think increased transparency is critical to promoting healthy public conversation on Twitter and earning trust." Twitter is making "data more public so that all can learn from it and we can also be held publicly accountable to it." Google has a "Transparency Report" (Google, 2021) that shares information

about content regulation, which says data for “content removal requests” is shared “in an effort to inform discussions about online content regulation.” By promoting transparency, companies can gain trust to avoid regulation.

Social media corporations will display warnings or information about misleading or false content without actually taking that content down. On fighting misinformation, Zuckerberg said (*Big Tech Company's Liability Shield*, 2020), “We've displayed warnings on more than 150 million pieces of content that have been debunked by our third-party fact-checkers.” YouTube displays an informational panel under videos and in search results that “are prone to misinformation” (YouTube Help, 2021). In 2020, Twitter (Roth and Pickles, 2020) introduced labels for “false or misleading content,” only taking misleading information down if it is marked as “severe.” These fixes allow for greater perceived trust without heavily impacting the experience of users.

Social media companies argue that regulation will inhibit free speech online. Zuckerberg (*Big Tech Company's Liability Shield*, 2020) argued that Section 230, which protects social media sites from being liable for users’ posts, “encourages free expression,” and without it, “platforms would likely censor more content to avoid legal risk and would be less likely to invest in technologies that enable people to express themselves in new ways.” In the same hearing, Google CEO Pichai maintained that Google’s “ability to provide access to a wide range of information is only possible because of existing legal frameworks, like Section 230.” Carl Szabo, vice president of NetChoice, a trade association comprised of members such as Google, Facebook, and Twitter, writes that required labeling of deep-fake posts would be a “slippery slope to the government deciding which speech is legal and which speech is illegal” (Szabo, 2020a). Bambauer, writing for the Brookings Institution, which is a non-profit supported by

Facebook and Twitter (2020), contends Section 230 reforms endangers free speech by making liability protection conditional on "difficult, if not impossible, decisions about contested concepts like truth and neutrality." This would mean that companies would need to take down posts to avoid "the threat of liability." By appealing to free speech, which is heavily valued in the U.S., companies can argue against regulation reforms.

Social media companies contend that regulation could incentivize less moderation of online platforms, leading to a less safe internet. Robert Winterton (2021), NetChoice's director of public affairs, argues that proposed legislation hampers online platforms' "ability to remove offensive content," preventing "kid-friendly content moderation." Szabo (2020b) argues that Section 230 reform would cause social media sites that apply "the type of content moderation required to protect children online" to be "liable for every user's post." This would disincentivize online platforms from any moderation at all, and lead to "an increase in hate speech, violence, conspiracy videos and other harmful content online."

Social media companies argue that regulation will stifle innovate and competition in online services. Bambauer (2020) contends that the "regulatory cost of these proposals could make it harder for start-up companies to compete with established internet companies," and worsen "antitrust and competition concerns." Szabo (2020a) argues that regulations will "increase the government's control over online platforms at the expense of free speech and innovation." Christopher Cox (*Online Censorship Reform,* 2020) argues that if "Section 230's clear limitation on liability" were to be reformed, it would be "difficult to imagine that most of the online services on which we rely every day would even exist in anything like their current form." And Winterton (2021) contends "legislation would significantly raise the costs of creating and running" online platforms, harming "the little guy" the most.

Criticism of Social Media

Some conservative politicians contend that social media algorithms are biased against them. Former president Trump claimed Twitter was "very discriminatory" against Republicans, saying "they don't treat me well as a Republican" (Trump, 2019). Republican senator Roger Wicker (*Big Tech Company's Liability Shield,* 2020), alleged that Facebook and Twitter unfairly blocked a story that revealed "communications between Hunter Biden and a Ukrainian official," while simultaneously allowing access to "the president's tax returns" and the Steele dossier. In the congressional hearing *Twitter: Transparency and Accountability*, Republican Senator Walden alleged "a member of this committee had her tweets and ads taken off the service because of a basic conservative message."

Some politicians contend that social media algorithms can be taken advantage of by foreign actors. Senator Jon Tester (*Big Tech Company's Liability Shield,* 2020) contends that Russians, "using a network of bots and fake accounts," spread disinformation. In the same hearing, senator Maria Cantwell alleged that Russians "masquerading as Americans" used "social media platform tools to interact and attempt to deceive 10s of millions of social media users in the United States."

Former tech insiders and journalists also argue that recommendation algorithms can become biased. Guillaume Chaslot, a former YouTube employee, ran a study that identified the most common YouTube recommendations for "Trump" and "Clinton," finding that in total, 86 percent favored Trump (Lewis, 2018). He is a part of the website Algotransparency, which says "Algorithms are gatekeepers of our information. Most content watched online is selected by an algorithm. Little is known about these algorithms… and they are not neutral" (Algotransparency, 2021). Reporter Kayleigh Donaldson (2019) alleged that top YouTube "search results for Katie

Bowman were videos dedicated to ‘proving’ she was a liar,” and that YouTube’s algorithm takes users from anything “vaguely feminist or female-centered” into a “rabbit hole of conspiratorial hatred.”

Advocates also worry that social media algorithms promote extreme content, or do not suppress it enough. According to the *New York Times*, a Facebook team “surveyed users about whether certain posts they had seen were “good for the world” or “bad for the world.” They found that high-reach posts — posts seen by many users — were more likely to be considered ‘bad for the world,’ a finding that some employees said alarmed them.” But when the team made an algorithm to reduce “bad for the world” posts, users opened Facebook less, so a less effective version was released instead (Roose et al., 2021). Tristan Harris (2020), president of the Center for Human Technology, and former Google employee, contends that YouTube’s recommendation system “does not know what’s valuable, ethical, or credible beyond what got the most clicks or watch time.” This causes YouTube’s algorithms to “recommended countless conspiracy theories.”

On creative platforms like Youtube, creators mistrust the recommendation algorithm for not promoting their content. Users will often comment saying “Commenting for the algorithm” on videos they hope will become more popular (Lucy, 2021). Brain Johnson argued in a video that the YouTube algorithm pushed creators to make more “clickbait-y” videos to be promoted, rather than making better content (Johnson, 2019).

Advocates contend that recommendation algorithms manipulate their users. According to Harris (2020), recommendation algorithms and the design of social media platforms exploit people’s reliance on “stopping cues,” giving users seemingly infinite content with no end or reason to quit. Sandy Parakilas (Orlowski, 2020), former operations manager at Facebook,

explains how social media companies constantly make small adjustments to their algorithms, so that “over time, by running these constant experiments, you develop the most optimal way to get users to do what you want them to do. It’s manipulation.” Shoshana Zuboff (Orlowski, 2020), member of the “Real Facebook Oversight Board,” contends that Facebook was “able to use subliminal cues to get people to go vote in the midterm elections.”

Advocates argue that social media algorithms isolate users. Pariser (2011) coined the term “filter bubbles” to describe how algorithms insulate users from opposing viewpoints. He contends that when Facebook’s algorithm noticed that he followed links from his liberal friends more that his conservative friends, it “edited them out,” so that conservative posts disappeared from his Facebook feed. Harris (2020) describes “personalized channels of automated news feeds” as “2.7 billion Truman shows,” alleging “social media has taken the shared narratives and facts that make society function” and split them among users.

Some advocates argue that social media companies and their algorithms are used as public services, and should be regulated as such. Harris (2020) compares social media platforms that “millions or billions of people rely on for the daily function of their lives” to public utilities, saying: “It is not a personal responsibility for people to protect themselves” from anything “which is inherently a commons or utility.” Therefore, the government should regulate social media platforms to “prevent societal harms.” Pariser (2019) contends that social media has “effectively taken control of our online public square,” and we should evaluate what “we need from them for the greater good.” Muldoon (2020) argues that large social media platforms should be regulated “to bring the service they offer back under public ownership and democratic control,” stopping them from maximizing “the amount of time individuals spend on the platform.”

Conclusion

Social media is similar to junk food or cigarettes. Though we know it is unhealthy, we continue to use it, and social media corporations are more than happy to keep us addicted. While companies can appear as if they are aligned with the public good and user's needs, in reality algorithmic fixes will always tend towards keeping users engaged, which can then be pointed to as "user satisfaction." Easy fixes such as content warnings and transparency over content moderation help somewhat, but do not change the issue of recommendation algorithms meant to exploit human weakness. Although new legislation and reform may make social media less enjoyable and harder to use, even cutting back on the scale of social media, this may be exactly what we need, just like how quitting cigarettes leaves room for healthier uses of time.

References

Algotransparency. (2021). Our manifesto. Algotransparency. https://algotransparency.org/our-manifesto.html

Amazon/Facebook/Google/Apple Dominance. 116th Cong. (2020).

Bakshy et al. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science* (June). 1130-1132. doi: 10.1126/science.aaa1160

Bambauer, D. (2020). How Section 230 reform endangers internet free speech. Brookings. https://www.brookings.edu/techstream/how-section-230-reform-endangers-internet-free-speech/

Big Tech Company's Liability Shield. 116th Cong. (2020).

Covington et al. (2016). Deep Neural Networks for YouTube Recommendations. *RecSys '16: Proceedings of the 10th ACM Conference on Recommender Systems* (Sep.). 191-198. doi: https://doi.org/10.1145/2959100.2959190

Cummings et al. (2005, February 23). Consumer acceptable risk: how cigarette companies have responded to accusations that their products are defective. *Tobacco Control* (Dec.). iv84–iv89. doi: 10.1136/tc.2004.009837

Facebook. (2019, February 6). No, Your News Feed Is Not Limited to Posts From 26 Friends. Facebook. https://about.fb.com/news/2019/02/inside-feed-facebook-26-friends-algorithm-myth

Fleming-Milici and Harris. (2020, August 13). Food marketing to children in the United States: Can industry voluntarily do the right thing for children's health? *Physiology & Behavior* (Dec.). doi: https://doi.org/10.1016/j.physbeh.2020.113139

Google. (2021). Google Transparency Report. https://transparencyreport.google.com/?hl=en

Johnson et al. (2019). Emergent dynamics of extremes in a population driven by common information sources and new social media algorithms. *Sci Rep* (Aug.). 1-9. Web of Science. doi: 10.1038/s41598-019-48412-w

Donaldson, K. (2019, May 4). YouTube’s algorithm is bad for women. SYFY. https://www.syfy.com/syfywire/youtubes-algorithm-is-bad-for-women

Harris, T. (2020, January 8). Unregulated Tech Mediation → Inevitable Online Deception → Societal Harm. Written Statement prepared for a Congressional Hearing. https://energycommerce.house.gov/sites/democrats.energycommerce.house.gov/files/documents/010820%20CPC%20Hearing%20Testimony_Harris.pdf

Lacy-Nichols et al. (2019, August 9). The politics of voluntary self-regulation: insights from the development and promotion of the Australian Beverages Council's Commitment. *Public Health Nutrition* (Feb.). *23(3)* 564-575. doi: https://doi.org/10.1017/S1368980019002003

Lewis and McCormick. (2018, February 2). How an ex-YouTube insider investigated its secret algorithm. *The Guardian*. https://www.theguardian.com/technology/2018/feb/02/youtube-algorithm-election-clinton-trump-guillaume-chaslot

Meyerson, E. (2012, August 10). YouTube Now: Why We Focus on Watch Time. YouTube. https://blog.youtube/news-and-events/youtube-now-why-we-focus-on-watch-time

Mosseri, A. (2018, January 11). Bringing People Closer Together. Facebook. https://about.fb.com/news/2018/01/news-feed-fyi-bringing-people-closer-together/

Muldoon, J. (2020, December). Don't Break Up Facebook — Make It a Public Utility. *Jacobin*. https://www.jacobinmag.com/2020/12/facebook-big-tech-antitrust-social-network-data

Online Censorship Reform. 116th Cong. (2020).

Orlowski, J. (Director) & Rhodes, L. (Producer). (2020, September 9). *The Social Dilemma* [Motion Picture]. United States: Exposure Labs, Argent Pictures, The Space Program

Pariser, E. (2011, March). *Beware Online "Filter Bubbles"* [Video]. TED Conferences. https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles

Pariser, E. (2019, July). *What Obligation Do Social Media Platforms Have to the Greater Good?* [Video]. TED Conferences. https://www.ted.com/talks/eli_pariser_what_obligation_do_social_media_platforms_have_to_the_greater_good

Perrin and Anderson. (2020, July 31). Share of U.S. Adults Using Social Media, Including Facebook, Is Mostly Unchanged since 2018. Pew Research Center. www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/.

Roth and Pickles. (2020, May 11). Updating our approach to misleading information. Twitter. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html

Szabo, C. (2020a, January 9). Jeopardizing Core U.S. Values: Tristan Harris's Misguided Regulatory Proposal for the Technology Industry. NetChoice. https://medium.com/netchoice/jeopardizing-core-u-s-values-and-endangering-the-internet-9e30b421f1a4

Szabo, C. (2020b, December 21). Why Section 230 protects kids, and what its critics get wrong. Protocol. https://www.protocol.com/section-230-protects-kids

Trump, D. (2019, April 23). Trump Twitter Archive. https://www.thetrumparchive.com/?searchbox=%22treat+republican%22

Twitter: Transparency and Accountability. 115th Cong. (2018).

Winterton, R. (2021, February 15). Florida governor's plan exposes children to more sexual content, extremism on internet. *Miami Herald.* https://www.miamiherald.com/opinion/op-ed/article249265330.html.

YouTube Help. Information panel giving topical context. (2021). YouTube Help. https://support.google.com/youtube/answer/9004474?hl=en

YouTube Team. (2020, August 25). Responsible policy enforcement during Covid-19. YouTube. https://blog.youtube/inside-youtube/responsible-policy-enforcement-during-covid-19/