Algorithms, Manipulation, and Autonomy

An Analysis of Facebook

Conor Brownell

Introduction

This story begins with emotional contagion, a scary-sounding psychological phenomenon that is experienced by anyone who is in the presence of other humans. Hatfield, Cacioppo, and Rapson define emotional contagion as "the tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of another person's and, consequently, to converge emotionally" (1993). Researchers like Fowler and Christakis extended this definition to study the spread of emotion over a large social network over time; they concluded via 20-year longitudinal analysis that happiness spreads amongst friends, spouses, and family members (2008).

Since 2008, the reality of large social networks has dramatically, irrevocably changed. By 2009, Facebook was by far the world's dominant social media platform, far outpacing the previous pacesetter MySpace (Arrington, 2009). In the process, Facebook broke ground on a new era of connectivity and redefined social networks and relationships forever. The early years of Facebook's dominance were characterized by rapid change driven by a thirst for more users, but in 2014, Mark Zuckerberg changed the company's motto from the (in)famous "move fast and break things" (Kelly, 2014). Unfortunately, in my opinion, their founding ethos—paired with a fundamental abdication of responsibility—resulted in a great many broken things, some of which were important to people's lives. This paper will examine Facebook's algorithms–which control what users are shown–and scrutinize what ethical responsibilities to its users Facebook may have as a company of its power and influence. These topics will be analyzed through the lens of a notorious research study.

The same year that Facebook changed its motto, it published a paper titled "Experimental evidence of massive-scale emotional contagion through social networks" (Kramer, Guillory, & Hancock, 2014). At the time of publication, Adam Kramer was a data scientist at Facebook while Jamie Guillory and Jeffrey Hancock were faculty at Cornell, a grouping that will be important. The study manipulated "the extent to which people were exposed to emotional expressions in their News Feed" and tested if this was correlated with "change in their own posting behavior" (Kramer, Guillory, & Hancock, 2014). Although the effect size of the study was miniscule—posting amount changed by a few tenths of a percent—the study had N = 689,003, so the effect was statistically significant. Although one can argue that such a small effect size is unimportant, the authors themselves counter by saying the effect would "[correspond] to hundreds of thousands of emotion expressions in status updates per day" (Kramer, Guillory, & Hancock, 2014).

This study raises many questions and resides in a zone of ethical murkiness. Furthermore, it became a hot button issue in the public eye. This is the study that launched a thousand think pieces and journal articles. In the course of this paper, I will sift through many of the perspectives on this issue and attempt to uncover a unified theory of Facebook's ethical obligations with regard to algorithms and user manipulation. It is not a dilemma of *whether* Facebook should use an algorithm, but of *how*.

Setting the Stage

Why did both the public and the research community react so strongly to this Facebook study? From a user perspective, people were not asked to participate in research and were unaware that they were doing so (Flick, 2016). Although users did sign a terms of service

agreement, they were functionally not aware that they could be involved in Facebook experiments (Flick, 2016). In essence, there was no informed consent. This was a subject that inflamed users and academics alike.

In addition, Facebook did not utilize outside ethical oversight even as they were attempting to manipulate users' emotions (Flick, 2016). This perceived lack of necessary oversight drew the ire of many academics who argued that such a study should be subject to an IRB-like review process (Boyd, 2014). Interestingly, the Cornell IRB was actually consulted, because two of the researchers were workers there. The IRB ruled that the study needed no review because it was not direct human research and the data was collected prior to the involvement of the Cornell researchers (Flick, 2016). This paper is not an analysis of IRB practices, but this ruling surely indicates prioritization of protecting the university as opposed to ruling solely on the basis of ethics. Those advocating for IRB-like reviews of corporate research need accept that board review processes are no panacea. Nonetheless, the point that this study should have undergone further review is a valid one.

The final—and perhaps most fundamental—issue taken with this study was that many users were *wholly unaware* that Facebook manipulated their feed *at all* (Flick, 2016). This is a crucial point and sets the stage for the rest of this paper. In many ways, this was a conflict of expectations and reality.

The other side of this issue would maintain that this study is merely business as usual. "At Facebook, virtually every change to the app, no matter how small or obviously beneficial, is thoroughly tested on different segments of the audience before it's rolled out to everyone" (Manjoo, 2017). A/B testing is the bedrock of Facebook's (and other platforms') efforts to

increase engagement (Boyd, 2014). So how is the emotional contagion study different from a best practice? This is essentially a discussion of business practices vs research.

Zooming Out

If we take a step back, much of the outrage can be traced to the context of the time into which this study arrived. In 2014, trust in Facebook and other social media sites was eroding; at that time the public trusted Facebook less than the IRS or NSA (Boyd, 2014). This study was published as the public was waking up to the "big data" era of online life. Users were primed to perceive Facebook in a negative, controlling light.

Now let us set aside for a moment the ethics of this study. Was any harm actually done? As discussed, the effect size of this study was incredibly small. However, one can argue—as the authors themselves did—that a small effect size multiplied by 689,003 people is meaningful. In my opinion, such a small effect size is of little consequence, especially since the study was performed for only a single week. Acknowledging this allows us to more deeply examine the bigger picture issues raised by this study, rather than getting bogged down in the immediate impact of a week-long experiment.

Algorithms: Can't Live with 'Em, Can't Live without 'Em

Social algorithms are not distinctive of Facebook, they are increasingly ubiquitous. David Lazer uses Google as a paradigmatic example of this rise; it was founded as "a simple deterministic ranking system" but has evolved to produce "[personalized] results on the basis of information about past searches and other contextual information, like location" (2015). Social algorithms are impossible to understand via their code alone: "the interplay of social algorithms and behaviors yields patterns that are fundamentally emergent" (Lazer, 2015). On the other

hand, people have an expectation that social media feeds are not manipulated. To users, social media and their telephone company are of a piece; they do not expect either to control how much contact they have with their social groups (Boyd, 2014). These challenges make the implementation of social algorithms a fraught situation. In a sense, users want to be catered to, without knowing that an algorithm is doing the catering.

Facebook is not alone in struggling with its implementation of social algorithms. For example, in 2016 Twitter replaced their reverse-chronological sorting with a new default algorithmic sorting method. After public pressure mounted, they gave users the option of picking their sorting method and even went the extra step of allowing users to block recommended tweets (Statt, 2018). Twitter seemingly realized that providing a straightforward chronological feed could be a market advantage as the public grew increasingly suspicious of algorithms and big data.

More problematic is YouTube's "Up Next" algorithm. The feature—which theoretically has the goal of keeping users engaged—can rapidly send users down a rabbit hole of "conspiracy videos, videos produced by hate groups, and pirated videos published by accounts that YouTube itself sometimes bans" following a search about a mainstream news topic (O'Donovan, Warzel, McDonald, Clifton, & Woolf, 2019).

Like Twitter, YouTube, and countless other platforms, Facebook algorithmically decides what to show people 24/7. These other examples serve to illustrate the challenges presented by social algorithms, but not all hope is lost. Users appreciate that the algorithm prioritizes things they like (Bucher, 2017). In my opinion, if Facebook were to switch to a chronological feed, users would be startled at how much refuse they had to sift through to find the things

they care about. In this sense, algorithms are certainly a useful tool that, as discussed, many people are unaware of.

However, sometimes the algorithm taketh away. Facebook can be a tremendous platform for fledgling artists to promote themselves, but doing so requires gaming the algorithm, with inconsistent success (Bucher, 2017). Users must fine tune posts to get maximum promotion by gaming the system. Sometimes the opposite problem occurs, and users are unable to control what they see. One user was repeatedly asked if she wanted to "poke" her ex-boyfriend, even though she had hidden all of his posts (Bucher, 2017). Users were also sometimes creeped out by targeted ads or posts clearly shown in reaction to their recent activity (Bucher, 2017). Algorithms run the risk of being too good at their job; users want to feel that they are in control of their social network not that their social media platform is determining what friends they interact with.

The crux of the algorithm question is that algorithms are updated and tweaked *constantly*. None of the aforementioned companies ask for your permission when they change how they recommend the next video or the next great follow. Twitter is an interesting case because they allow users to opt out, but that option does not undermine the effectiveness of their live event centric platform. Facebook, as mentioned, relies on filtering due to a high volume of content.

These algorithms are black boxes, so even if platforms did disclose every update the intended effects would be difficult to parse. Furthermore, the *intended* effects are potentially very different from the *actual* effects. The interplay of algorithm and human behavior is nigh impossible to predict.

To summarize, platforms do not disclose every update to their algorithms. Even if they did, the algorithms are complex and arcane, and modifications would be equally inscrutable. And even if their modifications were totally transparent, the algorithm's functional output is dynamic and unpredictable. Change, and therefore manipulation, is inherent to social algorithms.

Autonomy

The power dynamic at Facebook is increasingly asymmetrical as its user base continues to grow worldwide. If somebody found out a friend was manipulating them, they would likely end the friendship. But what can you do if the platform by which you communicate with your friends is doing the manipulating? Saying "find a new platform" is easier said than done when the largest tech companies are practically monopolies.

Technology made available partially makes our decisions for us, and this should be considered when undertaking ethically fraught experiments (Gertz, 2016). Facebook would claim that by agreeing to their terms of service users have agreed to participate in experiments such as the emotional contagion study. But part of Facebook's ethical calculus should include their asymmetrical power relationship with their users. Not only did Facebook not get informed consent—everyone, including Facebook, knows that nobody reads the ToS and are therefore not informed—but also users cannot autonomously decide to quit the platform. Their users are not fully independent in their decision to stay on Facebook. Their decision is partially made for them by Facebook's universality.

Conclusions

This is a fascinatingly complex issue. On one hand, Facebook's study violated the basic principle of informed consent and was not subject to exterior ethical review. On the other, it exemplified standard A/B testing that is a foundation of technological development. Industry members would likely argue "if there is no 'natural' News Feed, or search result or trending topic, what difference does it make if you experience A or B?" (Crawford, 2014).

In immediate terms, there was not a meaningful difference; the study essentially used a massive sample size to make marginal results significant. However, there is a problem with the big picture power dynamic of the user-platform relationship. Facebook is inescapable, and it controls the situation with its opaque algorithm. This muddies the ethical waters and takes us beyond a simple violation of informed consent. In my opinion, algorithmic manipulation was a tricky, insidious problem that Facebook inadvertently brought into the public consciousness with this study.

From the user perspective, this was a violation of expectations. In the aftermath of the study, Facebook apologized for the way the study was communicated, but not the study itself (Flick, 2016). This signals that they may have learned the wrong lesson. I fear that Facebook thinks that "the practice of algorithmic manipulation is acceptable but being transparent about the process is not" (Boyd, 2016). Did the reaction to this study teach Facebook that they should reduce transparency?

In the course of researching for this paper, I concluded that there is certainly much to be reformed. Although I am not persuaded that Facebook should simply have an IRB-like system, the mix of omnipresent algorithms with an inescapable, socially monopolistic platform is a

situation that necessitates change. Facebook cannot go on quietly manipulating its captive users. The status quo would be preferred to decreased transparency, but there are several outcomes that would be even better.

Future Ideas

Many people writing about this issue got hung up on the question of business operations vs research, but I do not find this to be a meaningful distinction. Tech companies such as Facebook have insignificant separation between research and product (Boyd, 2014). Quibbling over who precisely is manipulating your feed is losing the forest for the trees. With this in mind, all of the ideas presented here deal with Facebook as a whole.

One obvious modification that Facebook could make is to allow users to opt-in to A/B testing of different kinds (Crawford, 2014). It would be incredibly easy for Facebook to notify users of their new options. If users are not fully in control of their choice to be on Facebook, we should give them autonomy within the Facebook ecosystem. Users could choose between an algorithmically sorted feed (presumably the recommended choice) and a chronological one. They could choose to be a beta tester for different subsets of features. They could agree to try tweaks to the sorting algorithm. It is very possible to conduct sound research with a smaller sample than the near 700,000 that Facebook manipulated in the emotional contagion study, so reductions in participants is not a meaningful concern. Users could even be compensated for participating with early releases. In toto, the company should be responsible for *informing* the user what they are agreeing to, and the user should be responsible for *consenting* (Flick, 2016).

Within the algorithm itself, Facebook could provide options. To borrow an idea from artificial intelligence, Facebook could allow for different degrees of exploration and exploitation.

In my personal experience, I enjoy seeing posts from friends that I have not interacted with in a long time. It is enriching for me to be reminded of old friends and acquaintances. On the other hand, some people may want to remain "friends" with those people but do not want their updates to appear on their news feed. Allowing users to customize this would better address user needs and again would allow for increased autonomy within a restricted system.

These ideas—while useful and relatively easy to implement—do not address the larger questions raised here. Every platform that employs algorithms faces these same questions. Social media platforms and social algorithms are unregulated as of yet (although there is growing pressure to change that, including among some presidential candidates). Although I do not have an answer for what a big picture ethical solution would look like, I do know some requirements that I think it should fulfill. I see three major stakeholder groups: platforms, users, and institutions (not the least of which is our democracy). Platforms can be further broken down into business considerations and researchers.

As Boyd raises, such a solution should be conversational, and everyone on all sides should understand the other points of view (e.g. so that users see the value of A/B testing frameworks) (2014). These solutions must also be flexible and applied differently to different platforms. Ethics is contextual.

When I began this paper, I was faced with a dilemma not of *whether* algorithms should be used, but of *how*. I have suggested a few ideas that would, in my opinion, address some of the concerns with Facebook's current implementation, and I have laid out some requirements that solutions in general should adhere to.

References

Arrington, M. (2009). Facebook now nearly twice the size of MySpace worldwide. Retrieved from https://techcrunch.com/2009/01/22/facebook-now-nearly-twice-the-size-of-myspace-worldwide/

Boyd, D. (2016). Untangling research and practice: What Facebook's "emotional contagion" study teaches us. *Research Ethics*, *12*(1), 4-13.

Bucher, T. (2017). The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30-44.

Crawford, K. (2014). The test we can—and should—run on Facebook. *The Atlantic*, *2*.

Flick, C. (2016). Informed consent and the Facebook emotional manipulation study. *Research Ethics*, *12*(1), 14-28.

Fowler, J. H., & Christakis, N. A. (2008). Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ*, *337*, a2338.

Gertz, N. (2016). Autonomy online: Jacques Ellul and the Facebook emotional manipulation study. *Research Ethics*, *12*(1), 55-61.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current directions in psychological science*, *2*(3), 96-100.

Kelly, S. M. (2014). Facebook Changes Its 'Move Fast and Break Things' Motto. Retrieved from https://mashable.com/2014/04/30/facebooks-new-mantra-move-fast-with-stability/

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788-8790.

Lazer, D. (2015). The rise of the social algorithm. *Science*, *348*(6239), 1090-1091.

Manjoo, F. (2017). Can Facebook fix its own worst bug?. *New York Times Magazine*.

O'Donovan, C., Warzel, C., McDonald, L., Clifton, B., & Woolf, M. (2019). We Followed YouTube's Recommendation Algorithm Down The Rabbit Hole. Retrieved from https://www.buzzfeednews.com/article/carolineodonovan/down-youtubes-recommendation-rabbithole

Statt, N. (2018). Twitter will soon let you switch between chronological and fixed feeds. Retrieved from https://www.theverge.com/2018/9/17/17872276/twitter-algorithmic-timeline-settings-change-viral-tweet-response