Algorithmic Violence: An exploration of the YouTube recommender algorithm

Aditya Katakol

6833446

May 2020

MCMV17003

MA. New Media and Digital Culture

Supervised by Dr. Karin van Es

Utrecht University

Abstract

This article seeks to highlight the complicity of YouTube's recommender algorithm in promoting structural violence. It analyses the successors of the ElsaGate phenomenon, to identify the role of the algorithm in proliferating this violence, and embellishes the notion of algorithmic violence as a means to analyse this phenomenon. It develops a mixed methods framework guided by the principles of analytic autoethnography to draw correlations between the theoretical and practical functioning of the recommender algorithm, through the illustration of a case study of Minecraft Monster School. Aided by textual analysis and autoethnographic methods, algorithmic optimisation is identified as an agent of structural violence in reinforcing inequalities and hierarchies on the platform. It also pins accountability to algorithmic optimisation for being a vehicle of violence, for reinforcing violent content. The analysis suggests the role of the political economy of the platform in promoting sensational and divisive trends, and identifies the mechanisms through which it does so.

Words: 11240

Acknowledgements

I set out with an unanswerable research question as a means to study the ElsaGate phenomenon, guided by inquisitive spirit alone. With the aid of my supervisor Dr. Karin van Es and second reader Dr. Imar de Vries, I was able to develop some methodological structure with which to pursue this research. I am deeply grateful to my supervisor for the extended support in identifying relevant literature, development of research methods, continued assistance in developing the framework of the document and in extensively guiding my first attempt at formal academic writing. This research question would remain an inoperationalisable shower thought if not for her recommendations. I would also like to extend my thanks to Dr. de Vries for helping lay the foundation of research, and encouraging my pursuit of the research objective. Thirdly, I would like to thank my peers for lending a critical ear when I needed them. This project has been a culmination of all their efforts.

Table of Contents

1. Platform Culture

YouTube is a key player in the social media landscape, and is the most visited website domain besides Google as of 2019 (SimilarWeb data[1]). It is a video sharing platform that thrives off user-generated content watched and engaged with by other users. YouTube makes use of a recommendation algorithm that predicts content based on the user's interests, utilising metrics that remain still obfuscated to us. We do know from internal research that some criteria take priority in said promotion, like average watchtime of the content, clickthrough rate and user retentivity (Covington, Adams & Sargin, 2016, p. 194). However, numerous scholars (Bucher, 2018; Gillespie, 2017) have noticed a trend of content that is created with the intent of 'gaming' the recommender algorithm in order to get featured in algorithmically curated user feeds. As someone who grew up consuming YouTube content regularly and was moulded by its communities, I've noticed the emergence of certain patterns of content creation. This thesis sets out to interrogate how the YouTube algorithm contributes to a toxic culture through infrastructural violence.

The rate of uploaded content on YouTube currently averages around 450 hours a minute (Statista[2]), but has been growing rapidly over the last ten years. In 2010, programmer Chaslot helped develop the current recommendation system, in which it promoted content of "similar topics with a comparable sensibility" (Shaw & Bergen, 2019, para. 15), to help user engagement as well as revenue systems. The platform has reconfigured its algorithmic metrics numerous times over the years, but algorithmic moderation has not been able to adequately keep pace with the rate at which users are able to generate content. Over time, he too observed patterns of types of divisive and sensational content making it to the top – like flat-earth and holocaust denial content (Shaw & Bergen, 2019, para. 17). With the need to keep users engaged, we are left with an automated algorithm that "deals primarily with keywords, and is incapable of differentiating between 'Peppa goes to ballet lesson' and 'Peppa pulls gun on princess'" (Bridle, 2018, p. 229).

---

1. https://www.similarweb.com/website/youtube.com#overview
2. https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/

This idea was explored in James Bridle's '*Concurrency'* (2018), a piece of technocultural critique that brought attention to the phenomenon dubbed as ElsaGate. He describes a subecosystem on YouTube within which children's content is circulated, in which emphasis is put on hyperoptimising content with children-friendly tags and topics like "*Surprise Play Doh Eggs Peppa Pig Stamper Cars Pocoyo Minecraft Smurfs Kinder Play Doh Sparkle Brilho"* (Bridle, 2018, p. 219). He stipulates that in this platform culture, automation of content generation or following repetitive algorithmic formulae become inescapable, and as a result becomes "warped and stretched through algorithmic repetition and recombination" (Bridle, 2018, p. 225). The 'second-level' of these, as he describes, are disturbing and incorporate gory or sexual themes, but still succumb to the automated repetition of tropes. He describes these videos as feeding upon a system consciously intended to promote them for profit, but not accounting for the unconsciously generated emergent outcomes they may have (Bridle, 2018, p. 226). This algorithmic influence creates for an infrastructure enabling violence.

Rodgers and O'Neill's (2012, p. 402) model of 'infrastructural violence' argues that infrastructure makes for an ideal site of reflecting on the violence it participates in, through analysis of the suffering it affords. By analysing the stakeholders and functions, the design and access to the infrastructure, one is able to glean the cultural and political assumptions built into it (Rodgers & O'Neill, 2012, p. 406). This thesis seeks to model a similar heuristic to the virtual 'infrastructure' of the platform, to identify the role of the algorithm in perpetuating exploitative content. It seeks to illustrate that the recommender algorithm functionally plays a role beyond reflecting, but reinforcing violence by function of its agency. By providing a definition of algorithmic violence, the paper seeks to explore, 'How is the YouTube's recommender system complicit in proliferating violence?'.

This is a rather layered question, and necessitates the exploration of different subquestions to unpack the phenomenon in question. The first question: what is algorithmic violence? It involves a study of the political agency of algorithms, and their reinforcing functions. The second: What content does the recommender algorithm 'favour'? It regards the functioning of the recommender algorithm, and the ways in which that agency manifests. The final subquestion asks, how does this

acquired algorithmic favour go on to be an algorithmic vehicle of violence?

In a culture that is more influenced by algorithmic governance than ever before, critical media scholars have called for the interrogation of algorithmic systems, due to the emergent outcomes in their shadows (Striphas, 2015, p. 410). The emergence of this phenomenon has been a unique point of interest in technocultural discussion. Many have debated the potential repercussions on children, as easier and more malleable targets for the autoplay system (Buzzi, 2011, p. 244). Similar lines of discussion carry on regarding the radicalising power of the platform, whether users are guided to more extremist echo chambers as they engage (Tufekci, 2018, p. 2). But in this case of this immanent violence, Bridle goes on to identify "violence inherent in the combination of digital systems and capitalist incentives", as a further exploration into the political economy of the platform, enabling grotesque content (Bridle, 2017 para. 56). Without accountability, the infrastructure and revenue models of these ephemeral systems present multifaceted emergent outcomes that unchecked, can jeopardise social order. Similar studies show men receiving more ads for high-paying employment services (Datta & Tschantz, 2015, p. 92) or potentially being employed altogether (Dastin, 2019).

As such, this research is not a first foray into algorithmic ethnography and critique. It is preceded by research by critical media scholars, who have unearthed a great deal regarding algorithmic agency and mediation and how this shapes culture. The research concept was moulded by a chapter of Bucher's *'If...Then: Algorithmic Power and Politics'* (2018)*,* an investigation of the similarly unchallengeable Facebook timeline. However, much of the research regarding algorithms is inhibited by the algorithmic black box problem, in which researchers do not have the technical expertise required to interpret what is usually proprietary code subject to constant flux (Pasquale 2015, p. 6). Creative and ethnographic means have been employed to circumvent this problem, but do not employ analysis on the algorithm itself. While this is difficult for research, due to the personal nature of the reality presented by the algorithm, I believe there is insight in the tangible engagement with algorithms. To do so, I have developed a mixed methods framework, premised on the subquestions we sought to answer and based on Anderson's ideals of

analytical ethnography (Anderson, 2006). It makes use of textual analysis alongside an autoethnography of the recommender the algorithm to identify and corroborate identified patterns. I hope to achieve a more empirical, yet still exploratory, means to analyse how the YouTube algorithm helps stimulate a toxic platform culture. Before I can do that, I must develop the notion of algorithmic violence in my theoretical framework. This concept is critical in analysing the role of algorithmic agency in manufacturing toxic subcultures. In establishing this, I can proceed in operationalising this term in research of the YouTube recommender algorithm, to explore its role in the proliferation of violence.

2. [Algorithmic] Violence

In this theoretical chapter, I address the political nature of algorithms to illustrate the phenomenon of algorithmic agency; the ways in which the algorithm can express its political functions. Consequently, I outline how the algorithm influences content creation trends on the platform. Utilising the concepts of algorithmic recognisability and favour I then develop the notion of algorithmic violence. Finally, I address the concept of the 'algorithmic imaginary' (Bucher, 2019, p. 31) to create a pragmatic framework of analysis that bridges theory with practical research.

2.1 Politics and agency of algorithms

To demonstrate how the YouTube algorithm is responsible for perpetuating violence I must illustrate how the algorithm 'acts' to enforce its will. Traditionally, algorithms have been studied as "neutral" and "objective" processors of data. They weren't influences on culture themselves, but rather were identified as discrete objects that "may be located within cultural contexts or brought into conversation with cultural concerns" (Seaver, 2017, p. 4) Such an understanding fails to account for how algorithms are inherently political artefacts. They embody human discourse in their

knowledge logics, by function of what data is considered relevant, the logics by which it is processed, the predictions made by the algorithm, and so forth. Algorithms have influence over how they "enable and assign meaningfulness, managing how information is perceived by users, the "distribution of the sensible" (Langlois and Elmer, 2013 p. 94). Algorithms not only help us find information, they provide a means to know what there is to know and how to know it, and the means by which to participate in social and political discourse. As Seaver explains, "algorithms are not singular technical objects that enter into many different cultural interactions, but are rather unstable objects, culturally enacted by the practices people use to engage with them" (Seaver, 2017, p. 1).

Through automated processes algorithms assign relevance to media objects and, based on how this ranking works, in the end "attention is drawn to some things at the expense of others" (Just & Latzer, 2016, p. 250). The politics in question here are the "practices and capacities entailed in ordering the different ways of being in the world" (Bucher, 2018, p. 3). Tufekci (2015, p. 207) posits that algorithms are best conceptualised as actors in Latour's Actor-Network Theory - as non-alive computational agents that act with agency in the world. On the basis of how algorithmic recommender systems "sort, manipulate, analyse, predict" (Gillespie, 2010 p. 349), they inherently participate in political agency as well. It turns into a gatekeeper in its choices regarding visibility, in how it recommends, suggests, and provides users with what it has predicted to be the most relevant, hot, or interesting content to consume.

Algorithms, in a curatorial role, meld with actors including programmers and users and function as part of a "process of intermediation" for cultural products (Morris, 2015, p. 450). Similarly, Bucher (2018, p. 1) describes these platforms as "performative intermediaries that participate in shaping the world they purport to represent" by function of their algorithmic operations.It becomes important to realise this 'performative intermediary' function of algorithms in dictating the 'distribution of the sensible'. In looking at the ways that algorithms can enforce political capacities over others, we are able to identify a means for the algorithm can act with agency in this world.

2.2 'Algorithmically Recognisable'

From the critical perspective of political economy, the sociality and connectivity afforded by social media platforms and mediated through the algorithm exist as a function of a business model. As José van Dijck stipulates, "Social media are inevitably automated systems that engineer and manipulate connections" (2013, p. 28). Social media platforms 'do much more than just sell users' attention to advertisers: they actually help identify the very strategies through which attention can be fully harnessed' (Langlois & Elmer, 2013, p. 95). Seaver describes this function of the algorithm in an analysis of the captological turn. He observes how, instead of predicting explicit user ratings to guide recommendations, developers began to anticipate implicit ones. This set the tone for a captological approach to design, one in which the platform's ultimate goal is the creation of an inescapable technocultural landscape for its users, designed in benefit of its revenue model (Seaver, 2017, p. 4).

YouTube has similar captological interests in keeping the user engaged on the platform. It provides a foundational framework to understand the algorithmic logic employed by the platform for recommendation. In concordance to the revenue model, YouTube needs to attract viewers and keep them engaged by means of recommendations. Recommendations that are more likely to appeal to a user are better for engagement, and so the recommendation algorithm is optimised to promote certain metrics in content. Bucher also concludes that social media algorithms tend to only reward the 'right' kind of sharing, giving certain kinds of posts more visibility at the expense of others (Bucher, 2018, p. 36).

Users interested in having their content highly ranked in recommendations will tend to orient themselves toward these algorithmic systems, to make themselves "algorithmically recognizable" in the hopes that they will be amplified by them (Gillespie, 2014, p. 63). As Van Dijck's research shows in similar context, participants in cultural production are redirected to following algorithmically created trends, based on how social media contextualise different niches of everyday life (2013, p. 12). This algorithmic logic is a central tenet of platformization as "content

developers are progressively orienting their production and circulation strategies towards recommendation, ranking and other end-user facing algorithms of major platforms" (Nieborg & Poell, 2018 p. 4280). One such example is a case study in Bucher's analysis, in which a group of women "used their cleavage baring bodies as thumbnails to drive traffic" to their own videos on YouTube (2018, p. 128). As the YouTube algorithm rewarded only clicks at the time, these 'Reply girls' were able to obtain millions of views through their instrumental use of the algorithm.

The manifestations of the political agency of the algorithm are observable through the ways algorithmic favour is gained. Insights into the ways cultural production trends are shaped by algorithmic logics becomes a key site of understanding the means by which the algorithm can act in reinforcing social and cultural patterns. However, how do we gauge ourselves what the platform favours, before we can assume to draw connections between the political agency of the recommender algorithm and structural violence? How do we observe these algorithmically generated trends? The black-boxed nature of the recommender algorithm (Pasquale, 2015, p.2) calls for research gleaned from "a disparate array of sources in many different ways" (Seaver, 2017, p. 4), which may be insightful if the research emphasises on the cultures of the algorithm rather than specific configurations of it (Seaver, 2018, p. 6).

2.3 Imaginary

To critique the absolutism of the black box metaphor, Bucher (2019) proposes the study of "algorithmic imaginaries". She describes the algorithmic imaginary as users' perceptions about what the algorithm is and how it works shape their orientation towards it. While we cannot study the black box, we can begin to understand the performance of algorithms through the ways in which they are 'being articulated, experienced and contested in the public domain' (Bucher, 2019, p. 39). The situations in which people experience algorithms shape ways of thinking and discussing them. In analysing the affective and productive power of these imaginings, she posits

that we can see the influence of the imaginary on the functioning of the algorithm itself. She describes, "the social power of algorithms stems from the recursive 'force-relations' between people and algorithms", that give people a 'reason to react' to the algorithmic culture and its behaviour (Bucher, 2019, p. 39).

The algorithmic imaginary acts as the conceptual bridge between our assumptions about the political agency of the algorithm and how those assumptions result in certain practice. Based on the imaginaries circulating within the culture regarding what the algorithm favours, content creation patterns are funneled in directions encouraged by presumed algorithmic logic. The recommender algorithm promotes content concordant to its commercial aims, which encourages content creation optimised to the revenue model, like emphasising watch time or prioritising similarities to other successful videos (Bishop, 2018, p. 73). The captological interests of the platform contribute to its culture, causing particular types of content to rise to the surface. Bucher (2018, p. 129) corroborates that "..creators, in pursuit of view counts, adapt their practices to their understanding of the YouTube algorithms".

YouTube has fleshed out an extensive set of guidelines that supposedly provide content creators all the tools to enable their success on the platform, from branding to engagement – in the Content Creators' Playbook[3]. However, just like how Google releases limited information or advice on how to rank content, and actively discourages techniques used to 'game' its algorithm noting that there are "no secrets that'll automatically rank your site first in Google'; YouTube Creator Studio guidelines advise not to focus on algorithmic optimisation: "instead of worrying about what the algorithm 'likes,' it's better to focus on what your audience likes instead". However, the experience of content creators on the platform has been otherwise, and a significant amount of creator attention is directed at appeasing the recommender algorithm. We would need to identify how they adjust their practices according to how they think the algorithm works.

Bishop (2020, p. 1) provides a foundation for the study of "appeasing the algorithm" to achieve algorithmic favour. She posits that the algorithmic expertise provided by the 'experts' of the

3 YouTube's guide for content creators - www.thinkwithgoogle.com › youtube-playbook

platform regarding the metrics of the recommender algorithm is "a mix of data-informed assumptions that are weaved into a subjective narrative" (Bishop, 2020, p. 1). While they may present themselves as enemies of the algorithm, as they claim to reveal hidden algorithmic signals, experts' work often teaches creators to be complicit with YouTube and its business models. Bishop's notion of algorithmic lore is a study of how experiment data, theorization, and assumptions are weaved into a narrative on how algorithms work, and used as advice on how to successfully produce content. Her model is an interpretation of the algorithmic imaginary, in its understanding of how the subjective decision-making practices of human intermediaries continue to play a "significant role in even ostensibly algorithmic symbolic production" (Bishop, 2020, p. 1). Her reinforcement of the 'imaginary' enables the creation of a framework to interpret experts' opinions on algorithmic favour, to gain insight on the politics of visibility at play.

2.4 Towards algorithmic violence

In Bridle's *Concurrency* (2017, p. 236), he describes "a growing sense of something inhuman" in the proliferation of disturbing content by the YouTube autoplay algorithm. He identifies the architecture (of YouTube) as being 'hacked' to abuse children. He concludes, "[…] the structures we have built to sustain ourselves are being used against us […] in systematic and automated ways" . He describes this as an 'infrastructural violence done to all of us' beyond the existing violence in the content, as a byproduct of the political architecture of the platform. (Bridle, 2017, p. 236).

In seeking to understand this notion of 'infrastructural violence', I revisited the anthropological text authored by Rodgers and O'Neill (2012) on which it was modeled. Rodgers and O'Neill sought to reflect on infrastructure as a model for the analysis of societal violence, by analysing the infrastructural proliferation of suffering by nature of the cultural and political assumptions it espoused. Rather than attempting to capture material violence, it emphasises the study of broader, structural violence. By factoring the 'design and access (or lack thereof)'

(Rodgers & O'Neill, 2012, p. 404) to infrastructure, one could illuminate the political economy within which the infrastructure was located. They highlight examples of infrastructural manifestations of class struggle or racial prejudice in the architectural materiality of the city, which goes on to reinforce the status quo of unjust social conventions. The structural violence in question is substantially linked to issues of representation and visibility, over the hierarchy of political economy it seeks to fortify.

As we have just sought to unearth the cultural and political assumptions built into algorithms, the parallels to infrastructural violence are abundant. As the algorithm plays the role of the 'performative intermediary' (Bucher, 2018, p. 1) in distributing meaning, it enacts its agency over the visibility of content in many ways. Similar to the manifestations of social inequalities in the infrastructural interpretation, the algorithm reinforces the hegemonic dialogue of an algorithmically aware 'in-group' of content optimisers to various degrees, concordant to the hierarchies of its political economy. It seems appropriate to imagine the recommender algorithm as the virtual infrastructure of the platform, upon which reflection of the burgeoning violence in its communities becomes enabled. The 'design and access' metaphor is encoded in the politics of visibility that the recommender algorithm participates in, in the content that it favours by function of its 'algorithmic recognisability' (Gillespie, 2017, p. 63).

The notion of infrastructural violence emphasises "systemic forms of violence that occur through a society's effort to organise itself" (Rodgers & O'Neill, p. 411), implying the heterogeneity of societal stakeholders that contribute to structural violence. However, the distinctive power structure of the platform economy makes for a more convoluted site of research. 'Algorithmic favour' is obtained in a manner that is obscured to research by proprietary code, serving the revenue interests of the platform. This encourages algorithmic activation of relational impulses, similar to Bucher's research regarding Facebook's 'friendships' functioning as a suggestive force that encourages users to connect and engage with the people in ways that are afforded by and benefit the platform (Bucher, 2018, p. 5). She calls this engagement the 'programmed sociality' of the algorithm. Social media algorithms are also constantly in flux, modified by patterns of use and

engagement, which makes infrastructural reflection temporal and niche. They are "likely so dynamic that a snapshot of them would give us little chance of assessing their biases" (Pasquale, 2009, para 4).

The unique circumstances afforded by the platform economy and its hierarchies, as well as the personalised and mutable nature of the algorithm make the determination of a reflexive site for infrastructural violence impossible. To circumvent this, I propose a new heuristic of capturing infrastructural violence in concordance with the conditions of the algorithm; algorithmic violence. Bridle expresses his concern about the internet's way of amplifying latent human desires, in which it embodies visceral and violent identities. The algorithm goes on to proliferate this violence in its content as algorithmic structural violence; similarly reinforcing unjust social conventions through an intricate organizational network of metadata and algorithmic trends. In choosing to proffer visibility on certain content over others in the revenue interests of the platform, the algorithmic logic homogenizes content creation trends in patterns of favour. Algorithmic favour by function highlights certain genres of content creation and funnels content production trends towards them by enforcing inequality among content creators of various degrees of algorithmic compliance.

While I introduce the notion of algorithmic violence, it is by no means a new way of looking at the self-reinforcing capacity of the algorithm. There has been much discussion regarding the radicalising function of YouTube on malleable youth. As Tufekci (2018, p. 1) describes, "the YouTube's recommendation algorithm promotes, recommends and disseminates videos in a manner that appears to constantly up the stakes. Given its billion or so users, YouTube may be one of the most powerful radicalizing instruments of the 21st century". An excerpt from her column reads,

> A more likely explanation has to do with the nexus of artificial intelligence and
> Google's business model. For all its lofty rhetoric, Google is an advertising
> broker, selling our attention to companies that will pay for it. The longer people
> stay on YouTube, the more money Google makes.

What keeps people glued to YouTube? Its algorithm seems to have

concluded that people are drawn to content that is more extreme than what

they started with — or to incendiary content in general.

While she is discussing the specific tendency of the platform to pigeonhole users into consuming right wing content; which is still debated in the critical sphere (Ribeiro, 2019; Ledwich & Zaitsev, 2019); the idea is the same. As Seaver describes, "This study is in no way counterevidence to Tufekci's argument about radicalization, even though it shows a push to popularity rather than a push to fringe, because that is exactly what would be changed by personalization" (Madrigal, 2018, para 10). In enforcing its political agency favouring sensational and divisive content, the YouTube recommendation algorithm "allows for direct behavioural control of audiences by nudging them to certain kinds of consumption" (van Es, 2019, p. 235). While this control might not be a sinister motive but rather the entanglement of usage patterns and what Bridle describes as the platform's function of amplifying latent human tendencies, he does identify these manifestations to be tending towards violence (Bridle, 2017). Regardless of the intent, I only seek to highlight the exploitation and need for algorithmic accountability, through the purview of captology and violence.

3. Challenging the black box

This thesis employs a framework utilising a mixed methods approach. It seeks to utilise textual analysis and ethnographic methods to navigate the recommender algorithm of YouTube and study the culture of the platform. While this mixed methods approach provides for a rather contextual analysis, the ephemeral nature of the algorithm creates for personalised cultural realities that generate unique user experiences. This renders the recommendation algorithm unchallengeable by more traditional methods of empirical inquiry. However, I have spent many years immersed in the platform and its culture, and have some native knowledge of the discourse surrounding it. I believe this qualifies me with what Anderson describes as member status within

the culture of study. He posits that developing research methods in which the researcher's feelings and experiences are incorporated and their personal engagement is considered an important account for understanding the social world being observed, becomes key to illuminating the phenomenon at hand. While the imaginary does allow us to overcome the issue of how we think the algorithm 'acts', over how it 'acts', the YouTube algorithm is constantly in flux, making this research time-gated in validity. In developing a research form in which 'insider-knowledge' and subjective interpretation of experiences are an intrinsic part of the methods, we are able to research phenomena that afford "variable interpretations of events and cultural 'realities' " (Anderson, 2015, p. 381). However, these interpretations are subject to individual experience, and the personalised nature of the recommender algorithm leaves us uncertain with the inconsistency of user experience.

In the first section I explain how I intend to analyse the intent of YouTube's recommender algorithm through the heuristic of visibility, what it chooses to favour through a textual analysis of trends in the YouTube algorithmic imaginaries, providing us with guidelines of what is 'algorithmically recognisable' in practice. The second section observes a case study of a site of burgeoning violence, while observing its relationship with algorithmic optimisation. In the third section, we create a framework to observe the function of algorithmic favour as a vehicle for structural violence through the same case study, by manifesting as user recommendations in the platform.

3.1 Cultivating the imaginaries

We have earlier established in the framework that this algorithmic logic is a central tenet of platformization as content developers are progressively orienting their production and circulation strategies towards being recommended. Before we can make observations regarding 'algorithmic favour', it becomes a key ingredient in the study to identify how content seemingly achieves algorithmic favour, by function of its optimisation, by employing the algorithmic imaginary. The first

step in outlining algorithmic recognisability would be the identification of the imaginaries around which content creation trends are situated. The algorithmic imaginary is key for bridging the gap between the assumptions and the practice on the platform; by identifying the techniques utilised by platform experts in order to be successful by platformised standards, we are also identifying the means by which the platform is able to enact political agency by recommending certain types of content over others. "(Content creators) ideas about which types of content are made more visible by these algorithms influences choices about what to produce, when to circulate one's content and how to input metadata" (van Es, 2019 p. 235).

I developed a framework of analysing the imaginary through Bishop's heuristic of algorithmic lore. By analysing the most relevant videos regarding techniques to gain algorithmic success, they serve as algorithmically decanted success stories of complicity with YouTube and its business models as functions of being recommended. In the interests of cultivating the imaginary that influenced algorithmic optimisation trends, I sought to do a textual analysis of the texts made by the ten most influential algorithmic 'experts' giving advice on gaining YouTube popularity. The goal was to reveal how experiment data, theorization, and assumptions are weaved into a narrative on how algorithms work, and used as advice on how to successfully produce content (Bishop, 2020, p. 1). Through this analysis, we are able to gain insight into what is 'algorithmically recognisable', by function of the imaginary.

To establish my corpus I searched how to gain 'popularity', 'success' and 'views' on YouTube; on the platform of YouTube itself, and compiled the top results. This would create a pool of algorithmic lore, from which specific imaginaries circulating the platform could be gleaned. So further, in identifying the repetitions and commonalities across these texts in discussing the YouTube algorithm, we have generated a means to understand the imaginaries by which cultural trends are adapted to creators' understanding of the YouTube recommender algorithms and their agency. I identified ten sources of lore, eight of which were prominent algorithmic experts within the culture of YouTube, and I made use of two sources from outside the platform but culminated in the opinions and experiences of multiple content creators to provide insight. These sources were

[Brian G Johnson, Brian Dean, Jade Darmawangsa, Liam James Kay, Dan Lok, ThinkMedia, VideoInfluencers, TekkitRealm, HootSuite (blog) and Oberlo(blog)]. I compiled this list on the 27th of March from my residence, using my home internet network.

3.2 Reinforcing Violence

Following *'Concurrency'* (Bridle, et al), I sought out similar examples of algorithmically favoured content being promoted by the YouTube algorithm. Aiming to situate this project in more recent examples, I identified a case study to represent the research phenomenon, the genre of content featuring "Minecraft", a sandbox videogame targeted primarily at an adolescent demographic. AlgoTransparency, a project seeking to unravel the recommendation algorithm of YouTube, identifies Minecraft as a key link to children's content on the platform (AlgoTransparency, 2017[4]). Invoking my "member status" (Anderson, 2015, p. 379), I had prior discussed this interest with the moderators at /r/Elsagate - a discussion forum highlighting inappropriate content on YouTube. They had noticed an upswing of Minecraft related content carrying bizarre themes. Being natively involved on the subreddit, I identified an escalation pattern for this genre. "Minecraft Monster School" was one such case study.

Monster school was created by the YouTube channel Wilcraft Animations in 2012, and featured titular characters enacting the roles of students in a school. It was a fifteen episode series, following which the author moved on to different genres of content. However, following its success, many other content creators leapt onto the bandwagon of creating similar episodic content in the series and now is a popular content tag on the platform, bearing 16,37,000 pieces of content and 175% trend growth [monthly search volume] over the last 12 months (KeywordTool stats[5]). Minecraft is a videogame targeted at young children, and in Bridle's illustration of [structural] "violence inherent in the combination of digital systems and capitalist incentives" (2017, p. 236), he

---

4   https://algotransparency.org/youtube.html
5   Keywordtool.io/youtube search for 'Minecraft Monster School'

expresses apprehension towards "children ... being deliberately targeted with content which will traumatise and disturb them, via networks which are extremely vulnerable to exactly this form of abuse" (Bridle, 2017, p. 237). Through pattern identification of content espousing themes of violence, sexual situations, substances and toilet humour as is common in Elsagate, #monsterschool seems to represent a similar proliferation pattern as Bridle's 'Bad Baby' or "Finger Family". In these situations, the algorithm goes on to promote through recommendation and 'blow up' immanently violent content.

Having identified the pervasive visibility of the content category of 'Minecraft monster school' in the recommendation system, I believe it makes for an ideal case study for algorithmic violence. I have native experience with its espousal of violent-leaning content as established by 'insider-knowledge' (Anderson, 2015, p. 379), that is algorithmically reinforced as a vehicle of violence by function of its algorithmic favour. To validate my assumptions, I seek to analyse the algorithmic trends that guide the content creation in this genre, and identify patterns of algorithmic favour that encourage structural violence. In this section, I intend to suggest these genres tend towards algorithmic favour as a function of platform logic, but also act as sites of burgeoning violence, of what Bridle describes as sensational or divisive, for the algorithm to reinforce.

In order to render a corpus for analysis, I compiled the fifty 'most relevant' pieces of content belonging to the #monsterschool content tag over the last year, by inputting that metadata into YouTube's search function with those filters, and tabulating the data presented to me by algorithmic filtering. In performing a textual analysis of the metadata of this content, I would have some cognizance of algorithmic trends that guide its proliferation patterns, to identify how sites of violence are enabled by the algorithm. The videos were compiled on the 29th of March 2020 and tabulated, noting their titles, thumbnails and content tags. In identifying the trends that guide #monsterschool content, we can make correlations to structural violence, as well as broader categorisation into distinctive algorithmic phenomena; as guided by the YouTube imaginary identified in the first section. Then, I captured my observations regarding the tabulated data in analysis.

These two sections in conjunction offer insight into the theoretical functioning of the recommender algorithm in theory. The first demonstrates the 'intent' of the algorithm through the heuristic of the algorithmic logic that guides platformisation, the way algorithmic culture influences algorithmic trends in the logic that guides cultural production patterns favouring the promotion of certain types of content. The second analyses a case study as a site of violence to be reinforced by algorithmic optimisation, insinuating that it is enabled by algorithmic favour. Bearing an understanding of algorithmic violence and how the recommender algorithm can enact its 'complicity', we are left to identify the culpability of the algorithm in enforcing structural violence on the platform.

3.3 Floating to the top

In order to corroborate with my model of algorithmic violence, I must identify the tendency of the recommender algorithm to promote that content, and the patterns by which it does so. Seeking to explore if the YouTube algorithm tends to recommend violent content, I ethnographically navigated the platform in the interest of verifying such escalation patterns. I employed Anderson's techniques of analytic autoethnography (2015, p. 378) to develop a framework of analysis for that autoethnography.

Autoethnography is focused around the self and reveals, "personal investments, interpretations, and analyses. It embraces and foregrounds the researcher's subjectivity rather than attempting to limit it. He encourages developing research that accounts for the personal experiences of the researcher, to make use of 'insider knowledge' to study that which involves "variable interpretations of cultural 'realities" (2015, p 225). However, in remaining austerely committed to the research goal and verifying personal interpretations among others steeped in the culture, he hopes to remain more consistent to qualitative inquiry situated in traditional interactionism, hence the 'analytic' (p. 219). He encourages the use of assemblages of personal

accounts through texts, be they photographs, journals, recordings or otherwise, adapted to the research and incentive. Anderson warns that the researcher should allow for narrative visibility of the self in research, while still remaining analytically reflexive in the ethnographic practice about their position and involvement with the subject matter. I wish to be specific about the insights about violence reinforced by the algorithm in this section – not every content category is as readily or at all attuned to violence despite having algorithmic rabbitholes of their own. However, for the purposes of highlighting the ElsaGate phenomenon, I have identified a content tag from without the platform.

I sought to create a framework through which to isolate this favoured genre organically using Anderson's guidelines, and to identify algorithmic patterns in the recommendations. In categorising this information in an empirically sound manner, this section of the thesis attempted to illustrate the notion of 'algorithmic favour'. By pragmatically navigating the recommendations served up on autoplay in situ in a process of "inputting controlled data onto a platform and closely monitoring algorithmic output" (Kitchin, 2016, p. 24), I would have a tangible corpus upon which corporeal observations can be made. In accordance to the assumptions made by the thesis, by expressing my interest in Minecraft, I would be able to identify this algorithmically favoured content category, as well as gain insight as to the algorithmic trends that guide the recommender algorithm's visibility function.

I set out to be organically served autoplay recommendations from Minecraft Monster School, in order to identify the pervasive nature of algorithmically favoured content. If not, the ethnography would provide me some insights as to what the algorithm 'favours' in real time. To do so, I generated five disparate and prior unused YouTube accounts to use as eyes into the platform culture. These accounts attempt to be non-partisan in their relationship to the platform, and bear no cookie data. They each utilise a different network to access the platform, situated across Utrecht [not accounting for geographical factors having weightage on recommendations]. For the purposes of the analysis, each of these accounts had some interest in 'Minecraft', so I determined five arbitrary interests that a user might commonly pair with Minecraft, like tutorial, playthrough,

animation, song and funny. Through initial testing, I noticed the pervasive nature of the Monster School content tag as it was consistently recommended in but a few clicks. I use these observations to guide my method: navigating ten clicks deep into the autoplay for a particular search interest is still a valid site for observation of this phenomenon. In each 'click', I paid attention to observable patterns manifesting in the recommender algorithm, for visible algorithmic trends that may have become valent in the process. The corpus in question here are the recommendations of videos served in the autoplay. In doing so, I was able to obtain some insights into what specific content trends are made visible in this niche of cultural production, and reflect on the tabulated data and fieldnotes. If reasonable correlations can be drawn from the data, we can use the findings to answer the research question: *how* is the recommender algorithm complicit in algorithmic violence?

On the 29th of March across five remote networks, I set out to perform an autoethnography of five 'interests' in minecraft. I mapped each disparate account's served recommendations to their respective interests (looking at the top 5 recommendations on each video). Each time, I selected the next recommended video on the autoplay algorithm to succeed the served video to proceed with the mapping process. I then catalogued this information for interpretation of patterns that became valent in the process, which are reflected upon in the analysis. These catalogued recommendations are in Appendix 2, tabulated from the first click, mapping each recommendation over five clicks into an interest pool.

4. The mechanisms of algorithmic violence

4.1 Employing the Imaginary

This section of analysis instead emphasises the ways algorithmic logic and favour influences production trends through the ways that "content developers are progressively orienting their production and circulation strategies towards recommendation" (Nieborg & Poell, 2018, p. 4280). I described a framework by which I identified the commonalities among ten different sources of algorithmic lore; eight YouTube content creators who shared their theorization and experience, and two blog articles that make use of multiple case studies to illustrate the successes of their techniques. I chose these sources with reference to their visibility on the platform, and gleaned expert insights as to the functioning of the recommender algorithm. Following is a summation of the 'repetition and commonalities' found in the textual analysis in appendix 1 and the techniques advised by the experts to appeal to the algorithm.

The 'imaginary' in question is shaped by platform experts' understanding of how the algorithm works and what it does, and in essence, content creators on the platform have distilled their understanding of YouTube's recommendation metrics into four observable categories: (1) Videos of the same channel; (2) videos that are popular based on engagement, watchtime and views; (3) videos that a specific person might like based on their viewing history; and (4) videos with related metadata to trends. Only two of these are in the control of the creators [popularity based on engagement/watchtime/views, and videos bearing similar metadata to popular trends], much of the discussed techniques affiliate to one of these two categories. However, frequently both of these goals have similar pursuits, regarding identifying what 'trends' are popular and in what format to capitalise on that algorithmic favour.

Market research: Most experts recommend researching as to what kind of content may be in demand, whether it may be an unfulfilled niche or an already popular trend. VideoInfluencers, one of the algorithmic experts, recommends developing one's own brand of content and

maintaining consistency. They recognise the importance of creating a personal language for users to understand. Brian Johnson speaks about the niche category at his TEDx talk, discussing his success in creating guides and how-to videos for gaps that he thought he could fulfill. However, more experts recommend capitalisation on existing influential trends as the key to success. Brian Dean recommends "piggybacking off [of] other people's engagement" by following popular content ideas. He, among others, recommends the use of various third party APIs to monitor and track engagement, like KeywordTool, vidIQ and SocialBlade to boost engagement. These perform various functions, like statistically tracking trends and analytics for uploaded content for researching the self and other content creators. Collaborating with other content creators also provides for a way to mutually boost engagement.

'Similar' metadata: Brian Dean goes as far as to illustrate his 'sequel technique', as he lays emphasis on being in the 'recommended next video' as the secret to success. He encourages content creators to copy not only the metadata but also the content and style of existing successful content, to be the next video recommended after that one. He also emphasises the importance of keywords: saying the keywords in the video so that they algorithm can verify them as well as making use of the right ones, which involves copying from the target video page's source code. Some creators also discuss avoiding advertiser-unfriendly language, in order to actually be discoverable. The dichotomy of success between the 'bulls' and the 'bears' so to speak, remains absolute in this example as well, in identifying trends versus copying metadata. This notion of 'copycat content', of reuploading, remaking or compiling existing intellectual property could be a byproduct of this technique.

Creating engagement: Jade Darmawangsa believes that success is obtained from developing a consistent online identity, and encourages content creators to be "interesting to watch in your own way". OberloBlog is also keen on being compelling: advocating for regular schedules, encouraging subscription, creating playlists and linking to other social media. Jade specifies focus on engaging titles and thumbnails, and her recommendations involve a crisp introduction, having the title clearly illustrate the intent of the video, concise discourse and so on. **Thumbnails:** She

also has recommendations for the thumbnail, to be catchy, bright and clear. Her personal experience with thumbnails involves success when using two people's faces in them. Dan Lok reinforces this emphasis on engagement, illustrating ways specificity can be brought to creators' calls to action. He also calls our attention to the influence of titles. From Dan's experience, specific titles written in the style of classified ads - seeking to fulfill a need - gains engagement due to the curiosity drive it enables. We have discussed the importance of some types of metadata, but every expert describes at great length their techniques of manipulating metadata to gain engagement.

Self ranking: In engaging the audience with content, making use of the platform's engagement affordances becomes key. Content creators like ThinkMedia align their practices to maximise engagement according to the revenue model of the platform, by creating playlists, ranking content, collaborating with other creators, catering to certain demographics, and so on. **Maximising watchtime:** Experts also emphasise that watchtime is an important criterion for the recommendation metrics, encouraging longer videos with pattern interrupts to keep them engaging for longer durations. Since watchtime directly affects how much traction a video gets, having a hook or extended introduction are techniques that are offered.

Algorithmically compliant: Counterintuitive to Bishop's position, content creators seem rather self aware of their role in being algorithmically compliant. Much like the prior discussed 'monetisation-friendly keywords', there are numerous examples of how the platform has penetrated the culture. Many creators encourage the use of meta analytics by third party software in order to remain up to date with algorithmic trends. Some speak openly regarding their success as 'copycat' content creators, and discuss techniques to mirror metadata from other successful content. The metadata of content has, to a certain extent, been gamified as a part of the platform culture, behind the premise of creating 'searchable' content. Experts seem comfortable with the idea that their content creation patterns are modified in manners to suit the business interests of the platform, and acknowledge it without pause. They even reveal techniques by which they unearth hidden content tags in order to capitalise on trends.

Metadata loops: In paying close attention to the keywords and patterns of metadata applied to the content, experts seek to make their content immanently highly discoverable. By applying certain tags or phrases to the metadata or content of the video, it inherently receives algorithmic favour. Experts even recommend combining groups of trends in order to maximise engagement, as content on the platform tends to 'snowball' (appendix 1) in its discoverability, exponentially increasing as it succeeds.

There are many imaginaries that were more frequently visible in the analysis, like the importance of strong and engaging introductions or calls to action for the purposes of creating engagement. However, as the scope of research did not make content analysis feasible, I had to limit the aperture of the research to mostly metadata analysis. This goes for off-platform discipline as well, like cross-platform posting or frequency of uploads, which do not feature in this corpus as it only emphasises the most relevant content in the category, rather than emphasising the efforts of individual content creators.

In an entire subculture of algorithmic success, in which keyword analysis over popular trends and mimicking content creation patterns of other creators funnels content creation patterns in particular directions arbitrarily. The stripping away of integral identity to include all the relevant keywords in the title and to have two faces in a thumbnail encourage the conformity to an algorithmic standard. The entire machinery of content production begins to feel less organic as we come to similar conclusions that Bridle has in *'Concurrency'*, that the process *"*do[es] away with the human actors to create infinite reconfigurable versions of the same videos over and over again" (Bridle, 2018, p. 225).

4.2 Platform-enabled violence

Having developed some understanding of how the algorithm is able to act in ways that it is 'implicit' in influencing cultural production patterns, we now seek to identify how algorithmic recommendation can be a vehicle of violence by nature of the favour it grants certain types of content by examining the practice of content creators. To aid in the illustration of this phenomenon, I compiled the fifty most relevant #monsterschool videos over the last year, and performed a textual analysis of their titles and thumbnails to identify the ways in which the creators behind this content tag capitalise off of algorithmic recognisability. In this section, we make use of prior analyses to gauge the algorithmic favour of content in this content tag and seek to identify how algorithmic reinforcement can create niches of algorithmic favour that may act as sites of violence. In analysing the tabulated metadata of the videos, this section analyses how content creators in realtime engage in practices of algorithmic optimisation.

Using prior identified commonalities as a guide for 'favoured' content in the algorithm, we have created a framework for understanding algorithmic favour. However, due to the scope of limitations of the analysis, many metrics regarding the contents of the videos and schedules of engagement couldn't be used further in a framework of analysis only utilising few subsets of metadata. As I wished to observe the phenomenon broadly, I could not have pursued textual analysis of the videos to determine optimisation outside of scrapable metadata.

While discussing the patterns of metadata in this content, some specific trends become apparent. One such trend is identified by Bridle, where he describes the title structure as a "word salad to capture search results, sidebar placement, and "up next" autoplay rankings" (Bridle, 2018, p. 219). The titles contain multiple thematic angles that are hyphenated to capitalise on multiple trends, like *'Monster School – Herobrine Love Curse vs Girls – Sad Story Minecraft'.* It also becomes apparent that some of these trends are pastiches of trends outside of the platform, like manifesting in the format of popular mobile ads *('Monster School: Lv1 Crook vs Lv99 Boss (4) Minecraft animation'),* referencing other videogames (*'Monster School : Fortnite Dance Challenge*

3 - Minecraft Animation'), and other such virtual trends *('Monster School : SUPERHERO BOTTLE FLIP Challenge - Minecraft Animation').* As of recently, the coronavirus pandemic has been a recurring theme in #monsterschool animations ('*Monster School : LOCKDOWN ONE MONTH NO GOING OUTSIDE - Minecraft Animation').* Content creators capitalise off of existing popular 'trends', inside and outside the culture of the algorithm by being well-versed in them.

The videos also align with the broader imaginary regarding algorithmic favour. They are usually hovering very closely around the ten minute mark in length, and some feature several midroll ads set to multiple points during the videos. They make use of attention-grabbing 'clickbait' titles to capture viewers, like '*FAMILY VS FAMILY 2 - WHO IS THE STRONGEST MONSTERS - MONSTER SCHOOL'.* This seems to align with ideals of watchtime and hooks to optimise content. It becomes quite observable in our analysis as to the high degree of algorithmic recognisability that content in this bubble bears. It appears to satiate many algorithmic requirements as we compiled in our study of the imaginary, but we are still yet to reflect on the #monsterschool subculture as a site of violence.

Much of the content in the genre tends to follow similar patterns as enforced by both sections of the analysis, in maximising engagement with metadata management. The tag is rife with 'copycat' content, with many content creators uploading similar titles and thumbnails, riding each others' engagement to manufacture algorithmic favour in metadata loops. The videos seem to use the same characters as the original series, and feature a similar thematic language in which they present their thumbnails – with blender animations of titular characters featuring warped facial expressions in the thumbnail. I encountered some examples of content being featured on the list twice, which suggests that some content is reuploaded by other creators, engaging in theft of intellectual property. Content in this format also employs thumbnails in a visual language primed with markers from the imaginary. They are often bright and colourful, featuring two or more characters from the narrative. The characters are expressive and captured in a moment of conflict and a point of action. They make use of intergenre content in the thumbnails, by juxtaposing themes and objects from other media or formats into Blender animations of minecraft.

Content is also episodic or sequential, featuring multiple videos in the same narrative. This seems to align with the rules of ranking and playlisting content to maximise its engagement. Videos are presented in parts or ongoing series in order to keep an audience invested. They also feature similar metadata to one another outside of the thumbnails, bearing similar content tags and title structure to one another, like Brian Dean's sequel technique. As such, creators also repost their content, albeit slightly modified to suit new trends and practices, titled '*New'* or '*n#*', observably similar to Bridle's ideals of algorithmic recombination (2018, p. 225). Content creators are keen on employing the language of the algorithm in content creation patterns, and display awareness of the algorithm's sensibilities.

Content often features some form of contention or conflict between the characters involved, usually versatorially. It tends to promote competitive or combative situations and themes.The content tag also featured content that could be deemed inappropriate to the platform in a number of ways. Sometimes the content is brutal, featuring characters subject to varieties of torture like buzzsaws or being tied to railway tracks. Dead characters and graves seem to be a recurring theme in the thumbnails, sometimes with blood and armaments. Characters are visualised and described as psychotic or crazy, and bear traditionally horrifying facial expressions or postures. Incarceration and battle to the death are also recurring themes, along with various kinds of 'monsters'. Sexualised content is popular as well, featuring animated or juxtaposed pornographic characters. In a videogame where the characters are blocky and pixelated, many content creators take much effort to render sexualisations into their animations. Several videos feature attractive renditions of Mrs. Incredible from the Disney franchise in the monster school, for example. The 'love curse' content bubble of minecraft features thumbnails with multiple female minecrafters chasing or dominating the protagonists. The only female characters in the videos, if they appear, are to be subjects of sexualisation or gore - if not both at once. An alarming number of thumbnails display pregnant women being subject to various kinds of mutilations, and bondage seems to be a frequent theme as well. A particularly disturbing thumbnail of '*Epic Herobrine Challenge'* among others feature sexual assault in the thumbnails. Recommended content in the genre also delves

into scatological themes and 'toilet humour', which manifest usually in their thumbnails but not the titles as well. Frequently this is combined with coprophagic themes as well. The last of the prominent themes is horror, with the explicit intent to be frightening by combining other horror franchises with #monsterschool. Popular horror characters I noticed included Momo, Sadako, Baldi, etc have somehow become recurring characters in the Monster School genre, as a cross-media arc of the genre.

I shall not use this thesis in an exploration of violence in this content, instead choosing to emphasise the role of the algorithm in enabling and reinforcing it. In the production frenzy that it enables by influencing algorithmic trends, Bridle believes that automation of content generation or following repetitive algorithmic formulae become inescapable, and as a result becomes "warped and stretched through algorithmic repetition and recombination" (Bridle, 2018, p. 225). The structural violence in question here stipulates the notion of algorithmic favour, in promoting certain types of content that are more concordant with the political intent of the algorithm. By commercially and in terms of visibility, enabling content the algorithm exerts a mechanism of control that Bishop describes as 'the violence which is exercised upon a social agent with his or her complicity' (2018, p. 71). Her research regarding the political aim of recommendations encourages questions regarding the intent of algorithmic favour in case studies of violence.

In seeking to analyse the case study of *'Minecraft Monster School'* as a site of violent content that also espouses a high degree of algorithmic favour, we are able to illuminate the correlation as manifesting in the algorithmic reinforcement of violence as an emergent outcome. By analysing the reinforcing function of the algorithm, by function of the notion of algorithmic favour, we can go on to illustrate how the algorithm participates in propagating structural violence.

4.3 Filter bubbles

The purposes of this section of analysis is to advocate practically the role of the algorithm in promoting certain types of content over others, by identifying 'recognisable' or 'favoured' content promoted by the algorithm. I sought to interact with the algorithm across various interest groups still related to minecraft, in effort to identify observable patterns for content creation that are algorithmically favoured. To do so, I operationalised a theoretical framework centred around the case study of #monsterschool, and sought to identify tangibly how "particular types of content rise to the surface" (van Es, 2019, p. 234) by identifying that which is promoted by the autoplay system. Over the course of my interactions with the YouTube recommender algorithm across five independent and disparately generated YouTube accounts bearing similar interests in the children's videogame of Minecraft, I was able to encounter these recommendations in an organic manner. From initial hypothesis, the pervasive nature of this content genre made it readily discoverable for analysis – and I was able to replicate these results. Across five disparate accounts labeled A-E, four out of five accounts were promoted the #monsterschool content tag. Across this autoethnographic navigation of the recommender algorithm, I would also keep an ear to the ground regarding organically observable patterns in the content and metadata.

The first such observable trend of note is that content creation tends to be promoted in bubbles. We already refer to Eli Parisier's filter bubbles[6], the isolation of interests of an individual by algorithmic personalisation such that they are trapped in an ideological bubble. This notion is rather similar, with keyword-optimised content being served together in 'bubbles' profiled by algorithmic personalisation. Account B identified one such bubble trend as *"Noob Vs Pro Mining in Minecraft"*, and Account C identified others in *"Battle Royale"*, *"Illager vs Pillager"*, *"Derp Infection"* and *"Monster School"*. When looking into minecraft related content, entering one of these 'algorithmically favoured' content loops encourages recommendations to be served from within the same metadata and categories. Many of these videos were sequential, but most of the recommendations were from other channels that make 'copycat' content or compilations of existing

---

6   Beware online "filter bubbles" | Eli Pariser Ted Talk https://youtu.be/B8ofWFx525s

content, with the intention of 'piggybacking' (Brian Dean, appendix 1) off other creators' engagement. Several of these pieces, as I identified in the comments, were reuploads of existing intellectual property, only adding to this notion. Regardless, the metadata creates inescapable bubbles if one chooses to utilise only autoplay to guide their recommendations. The idea of getting 'trapped' in these 'filter bubbles' aligns with Seaver's captological heuristic of looking at algorithms. He posits that they exert an effort to create an inescapable technocultural landscape for its users designed in benefit of its revenue model (Seaver, 2017, p. 428).

The concept of 'filter bubbles' lies within the threshold of similarity to that of an echo chamber – highly curated content intended to enact confirmation bias by reinforcing particular political opinions over others. The algorithm seemingly participates in enabling this bias by lending favour to metadata-optimised content by function of its metrics. While this notion is explored in greater detail in the final section of this analysis, it does become important to note how this metadata-oriented recommendation leads back to content bubbles. Account E, interested in watching minecraft playthroughs, was recommended a popular guide by the creator 'Python's World'. The same author also published a series of content entitled, *"Lucky Minecraft"*, *"Satisfying Minecraft"*, *"Cursed Minecraft"*, and so on. The latter was following what UrbanDictionary describes as "A type of internet trend that includes images that are abnormal or off-putting", but linked to a video entitled *"Love Curse Part 2 Minecraft Monster School"* featuring a more literal curse.

However, it is this notion of captological bubbles that aids in the operationalisation of my research concept; *"Minecraft Monster School"* is one such applicable filter bubble and makes for an ideal site for the reflection on the notions of algorithmic favour as well as sites of percolating violence. Across multiple accounts and despite having variations in search interests and formats, I was redirected to the same series or sequences of content, like Prison Escape or Monster School from 'minecraft funny'. From each account, it was only a matter of a few recommendations (3-6 for four accounts) to be coaxed into these rabbitholes of metadata-optimisation.

4.4 [Infra]Structural Violence

We set out to identify the infrastructural violence that the algorithm is complicit in. Speculating that the algorithm is a source of suffering and reinforcement of cultural inequalities, we identified a framework through which we could eke out an understanding of algorithmically enacted violence. In identifying how algorithmic violence manifests, how the algorithm can enact its agency in influencing cultural production trends and thus how it can be complicit in proliferating violence, we have utilised this framework to illuminate the phenomenon of algorithmic violence.

The algorithm exerts its political agency in the interests of its revenue model, to harness maximum viewer attention while selling advertiser-friendly content, through the metrics of its recommendation. It does so by promoting the visibility of certain genres of content, those which are compliant with its implicit metrics. Content creators optimise their content towards algorithmic trends, motivated by platform logics, and generate filter bubbles of homegenic content guided by algorithmic favour. The hierarchies imposed by the recommender algorithm on content creation patterns bearing degrees of compliance to algorithmic trends represents a structurally violent mechanism that can proceed to reinforce the status quo, which may sometimes represent violence similar to Rodgers and O'Neill's unjust social conditions (2012, p. 404). This mechanism is representative of algorithmic violence.

Due to the massification of the content, themes are explored in various heuristics, and sometimes the violent ones stick. As Bridle identifies, in a rush "to accumulate ad revenue, [they] are feeding upon a system which was consciously intended to show videos to children for profit. (2018, p. 236) The unconsciously-generated, emergent outcomes of that are all over the place. He addresses the violence too,

> ...the level of horror and violence on display. Some of the times it's troll-y gross-
> out stuff; most of the time it seems deeper, and more unconscious than that. The

internet has a way of amplifying and enabling many of our latent desires; in fact, it's

what it seems to do best. I spend a lot of time arguing for this tendency, with

regards to human sexual freedom, individual identity, and other issues. Here, and

overwhelmingly it sometimes feels, that tendency is itself a violent and destructive

one.

In discussions about the political economy of the platform, Bishop finds similarly that the

recommendation algorithm goes on to enforce hierarchies of visibility in vloggers, in degrees of

algorithmic compliance. She identifies the 'compliance' to be linked to the hegemonic messaging

of feminine identity, in correlation with their marketability towards cosmetic products. (2018, p. 69).

This research is also an exploration of algorithmic violence by the platform's use of agency, but

identifies the 'complicity' of the algorithm in proliferating that algorithmic violence as well. However,

her research does not explore the phenomenon from the perspective of violence as I am

incentivised to do. In my case study, the algorithm feeds off sites of violence, enabling content

creation in violent patterns to exploit the clickability of 'sensational or divisive content'. But why?

Bridle identifies the violence to be a byproduct of human desires percolating into the

algorithmic system, which manifests in the form of ElsaGate-like content (2017, p. 237). The

spread of the violence could also be attributed to similar arguments that Tufekci (2018) makes

when she discusses the polarising function of the platform. She posits that the algorithm constantly

'ups the stakes' due to the interests of the platform in keeping you increasingly interested. There is

still much ambiguity surrounding the incentives behind ElsaGate content, as there are always

exceptions to any generalisations. One currently popular opinion I've seen frequently on the

subreddit involves paedophilic grooming of children, but there's little evidence to suggest that is

true. Regardless, the intent of the research was to identify the structural mechanism by which the

algorithm can participate in reinforcing violence, as a vehicle of violent content.

5. Algorithmic accountability

I sought to demonstrate the notion of algorithmic violence as an algorithmic function of reinforcing hierarchy on the platform, in response to lack of accountability the platform exhibits regarding moderation of this content. Bridle ends his essay writing, "Responsibility for its outcomes is impossible to assign but the damage is very, very real indeed". In an attempt to challenge the impenetrable YouTube algorithm, I sought out to identify the emergent shadows in the cultural forms of the algorithm, as Striphas (2010, para 3) encouraged. In terms of future research, the concerns at hand are twofold.

In the interest of the ElsaGate phenomenon, I wish there were more feasible methods to identify the impetus of the whole thing – I suspect that it is made more difficult due to multiple sets of actors with different intents, but we lack the criterion by which to distinguish. There is limited research or knowledge otherwise about this phenomenon; I have seen a networked pattern analysis of its proliferation to aid in its detection, but little about the origin or reason despite blind ideals of capitalist exploitation. However, in the interest of algorithmic violence, I believe that there are a innumerable applications to interrogate the ways that cultures of algorithms manifest in our social, cultural and political lives. It is quite similar to Bucher's notion of programmed sociality (2018, p. 4) in identifying the ways algorithms activate relational impulses as intrinsic functions of their captological models.

We live in an era where algorithmic governance and prediction is inescapable, and the algorithms have ways of deeply moulding our attitudes, beliefs and values. They inherently are political objects that have a vested interest in involving the user within their politics. Tufekci and Bucher discuss the influence the algorithm can have on election results or the nature of one's friendships. The percolation of violence into a system that drip feeds entertainment and education to children especially can have potentially dangerous consequences. The case study only presents a narrow view of the implications of the phenomenon, and we should remain critical of the emergent implications of such algorithmic cultures. The minecraft example was limited to blender

36

animations in its representations, but another popular example, Toy Freaks, features a father and daughters enacting some of the algorithmically oriented trends, which sometimes includes videos of the children vomiting and in pain. Other frequently noted examples include algorithmically produced content, as they use easily available character models and motion capture libraries to endlessly and meaninglessly generate content by some machine metrics. How does this affect us culturally? How does it affect our children?

I also want to use this section to reflect on the assumptions I have leaned on and the methodological framework I employed. The phenomenon in question of course is highly subject to variance and lacks reproducibility. I have attempted to document my findings as objectively as I could, but I went into this research bearing the objective of illuminating this phenomenon through my methods. While I utilised pseudo-personas in navigating the recommender algorithm, it was still subject to my decision making regarding criteria and interpretations. However, Anderson predicts that autoethnographers should expect to be involved in the construction of meaning and values in the social worlds they investigate. This is an expected challenge in the investigation of highly personalised and algorithmically curated phenomena. There are also concerns regarding the assumptions of the functioning made by the recommender algorithm, which I have attempted to circumvent with the notion of the algorithmic imaginary – the dichotomy between what an algorithm does and what users think it does. However, at the end of the day, much of the creative research methods regarding how the algorithm functions must resort to some sort of 'reverse engineering' (Kitchin, 2016 p. 15) and hoping to draw deterministic correlations. Thirdly, the consistency of the algorithm's 'actions' remain uncertain due to the ephemeral and mutable nature of the algorithm, which is constantly modified by internal testing. YouTube has even made great efforts to improve moderation, but it is still (or suspiciously) lacking.

The lack of reproducibility of the case study also attracts questions regarding the pervasiveness of violence on the platform. However, I hope to illustrate, even by drawing reaching correlations, that this is a real problem on the platform. I was able to quite consistently discover the

filter bubble of Minecraft Monster School across multiple accounts, multiple networks and multiple

locations. The phenomenon also manifests in various other ways and bubbles, as many content

creators in the imaginary analysis encourage, and creates a banal homogeneity in content

creation. Enough of these bubbles exist that I have observed members of my family innocuously

wind up in bizarre communities through no intention of their own.

Belonging to a country rife with political manipulation of new media services to radicalise,

tribalise or incite voter responses, it is paramount to be critical of the algorithmic cultures that

encircle us. I have witnessed our cultural tics being manipulated in the interest of political gain, and

I worry for the exploitation we might suffer while lacking the knowledge of how to interrogate the

ephemeral and proprietary code of algorithms. However, while these worries are sensible in the

current landscape of data jurisdiction, I'd like to summarise the perspectives of the thesis through

an argument by Tufekci (2018, p. 5). "This state of affairs is unacceptable but not inevitable. There

is no reason to let a company make so much money while potentially helping to radicalize billions

of people, reaping the financial benefits while asking society to bear so many of the costs."

Works Cited

Ananny, M. (2011, April 14) "The Curious Connection Between Apps for Gay Men and Sex

     Offenders." *The Atlantic.* Retrieved from

     http://www.theatlantic.com/technology/archive/2011/04/the-curious-connection-between-

     apps-for-gay-men-and-sex-offenders/237340/

Anderson, L. (2006). Analytic Autoethnography. *Journal of Contemporary Ethnography,* 35(4),

     373–395, https://doi.org/10.1177/0891241605280449

Bishop, S. (2020). Algorithmic Experts: Selling Algorithmic Lore on YouTube. *Social Media +*

     Society. 6(1), 1-11, https://doi.org/10.1177/2056305119897323.

Bishop, S. (2018). Anxiety, panic and self-optimization. *Convergence: The International Journal of*

     Research into New Media Technologies, *24*(1), 69–84.

     https://doi.org/10.1177/1354856517736978

Bridle, J. (2018) Concurrency. *New Dark Age: Technology and the End of the Future,* 216-238

     London: Verso Books

Bridle, J. (2017, November 6). Something is wrong on the internet. *Medium.* Retrieved from

     https://medium.com/@jamesbridle/something-is-wrong-on-the-internet-c39c471271d2

Bucher, T. (2018). *If... then: algorithmic power and politics*. New York: Oxford University Press.

Bucher, T. (2019). The algorithmic imaginary: exploring the ordinary affects of Facebook

     algorithms. *The Social Power of Algorithms*, 30–44.

     https://doi.org/10.4324/9781351200677-3

Buzzi, M. (2011). What are your children watching on YouTube?. *Advances in new technologies*

     interactive interferences and communicability. 243-252 Berlin, Germany:

Covington, P, Adams, J & Sargin, E. (2016). Deep Neural Networks for YouTube

     Recommendations, *RecSys* (*ACM)* Boston, MA, 191-198

Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against

     women. *Reuters.* Retrieved from

https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

Datta, A., Tschantz, M., & Datta, A. (2015). Automated Experiments on Ad Privacy Settings. *Proceedings on Privacy Enhancing Technologies* (1), 92–112. https://doi.org/10.1515/popets-2015-0007

Es, K. V. (2019). YouTube's Operational Logic: "The View" as Pervasive Category. *Television & New Media*, *21*(3), 223–239. https://doi.org/10.1177/1527476418818986

Gillespie, T. (2015). Algorithmically Recognizable: Santorum's Google Problem and Google's Santorum Problem. *Information, Communication & Society.* 20, 63-80. https://doi-org/10.1080/1369118X.2016.1199721

Gillespie, T. (2012). Can an Algorithm Be Wrong? *Crowds and Clouds* Limn. 2. Retrieved from https://escholarship.org/uc/item/0jk9k4hj

Gillespie, T. (2010). The Politics of 'Platforms'. *New Media & Society.* 12, 347-364.

Gillespie, T. (2016). #trendingistrending: When Algorithms Become Culture. *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies.* 52-76. Routledge, London.

Gillespie, T. (2014). The Relevance of Algorithms. *Media Technologies: Essays on Communication, Materiality, and Society*, 167-195. Cambridge (Mass.): MIT Press.

Just, N., & Latzer, M. (2016). Governance by algorithms: reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, *39*(2), 238–258. https://doi.org/10.1177/0163443716643157

Kitchin, R. (2016).Thinking critically about and researching algorithms. *Information, Communication & Society* 20(1): 14–29.

Langlois, G. (2013). Participatory Culture and the New Governance of Communication: The Paradox of Participatory Media. *Television & New Media*, *14*(2), 91–105.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. 1-14 Oxford: Oxford University Press.

Ledwich, M., & Zaitsev, A. (2020). Algorithmic extremism: Examining YouTube's rabbit hole of

radicalization. *First Monday.* https://doi.org/10.5210/fm.v25i3.10419

Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of

apps. *New Media & Society, 20*(3), 881–900.

Lupton, D. (2016). The diverse domains of quantified selves: self-tracking modes and

dataveillance, *Economy and Society*, 45(1), 101-122

https://doi.org/10.1080/03085147.2016.1143726

Maack, M. (2019, June 14). 'YouTube recommendations are toxic,' says dev who worked on the

algorithm. *TheNextWeb.* Retrieved from

https://thenextweb.com/google/2019/06/14/youtube-recommendations-toxic-algorithm-

google-ai/

Madrigal, A. (2018, November 8). How YouTube's Algorithm Really Works. *The Atlantic.* Retrieved

from https://www.theatlantic.com/technology/archive/2018/11/how-youtubes-algorithm-

really-works/575212/

Morris, J. (2015). Curation by code: Infomediaries and the data mining of taste. *European

Journal of Cultural Studies*, *18*(4-5), 446–463. https://doi.org/10.1177/1367549415577387

Nieborg, D., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent

cultural commodity. *New Media & Society*, *20*(11), 4275–4292.

https://doi.org/10.1177/1461444818769694

Pasquale, F. (2009, November 18). Assessing algorithmic authority. *Madisonian.net.* Retrieved

from http://madisonian.net/2009/11/18/assessing-algorithmic-authority/.

Pasquale, F. (2015). Introduction. *The Black Box Society. The Secret Algorithms That Control

Money and Information*. 1-41 Cambridge: Harvard University Press.

Ribeiro, M., Ottoni, R., West, R., Almeida, V., & Meira, W. (2019). Auditing radicalization pathways

on YouTube. *Proceedings of the 2020 Conference on Fairness, Accountability, and

Transparency*. 131-141. https://doi.org/10.1145/3351095.3372879

Rodgers, D., & O'Neill, B. (2012). Infrastructural violence: Introduction to the special issue.

Ethnography, *13*(4), 401–412. https://doi-org/10.1177/1466138111435738

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems.

Big Data & Society. 1-12. https://doi-org/10.1177/2053951717738104

Seaver, N. (2019). Captivating algorithms: Recommender systems as traps. *Journal of Material*

Culture, *24*(4), 421–436. https://doi-org/10.1177/1359183518820366

Seaver, N. (2013). Knowing Algorithms *Media in Transition 8,* Cambridge, MA, MIT Press

Shaw, L. & Bergen, M. (2018, April 26). YouTube's Plan to Clean Up the Mess That Made It Rich.

Bloomberg. Retrieved from https://www.bloomberg.com/news/features/2018-04-26/youtube-

may-be-a-horror-show-but-no-one-can-stop-watching

Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, *18*(4–5), 395–412.

https://doi-org/10.1177/1367549415577392

Striphas, T. (2010, July 13). How to have culture in an algorithmic age. *Differences and*

Repetitions, Retrieved from http://www.diffandrep.org/2010/07/13/how-to-have-culture-in-

an-algorithmic-age/.

Tufekci, Z. (2015). Algorithmic Harms beyond Facebook and Google: Emergent Challenges of

Computational Agency. *Colorado Technology Law Journal*. V13, 203-218

Tufekci, Z. (2018, March 10). YouTube, the Great Radicaliser, *New York Times.* Retrieved from

https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html

Van Dijck, J. (2013). Engineering Sociality in a Culture of Connectivity. *The Culture of Connectivity:*

A Critical History of Social Media. Oxford: Oxford University Press, 1-41