Recommended for You

Exploring user perceptions of the YouTube recommender system

Student Name: Cindy Oscar Moore
Student Number: 505527

Supervisor: Dr. João Ferreira Gonçalves

Digitalisation, Surveillance and Societies
Erasmus School of History, Culture and Communication
Erasmus University Rotterdam

Master Thesis
June 2022

Recommended for You

ABSTRACT

YouTube is the second-most used social media platform worldwide and its popularity is partly due to its refined recommender system. This system provides users with specifically tailored video suggestions and it is based on an algorithm that is responsible for over 70% of the daily watchtime on the platform. The best-watched genre on YouTube concerns entertainment videos and the objective of this research is to inquire about user perceptions of the recommender system and how this affects entertainment content consumption. The study at hand provides a literature review on recommender systems, algorithmic imaginary, user consumption behavior and the uses and gratifications theory in light of the YouTube recommendation algorithm. However, literature that tied together user perceptions of the recommendation algorithm to YouTube entertainment content consumption was still lacking. Therefore, the research question of this study is as follows: *To what extent do perceptions of YouTube's recommendation algorithm shape user entertainment content consumption?*

More specifically, it was studied whether user perceptions of the recommendation algorithm affect perceived entertainment content diversity, watchtime, recommendation satisfaction and perceived user agency. Moreover, it was inquired about whether content, social, process and/or technology gratifications influence recommendation satisfaction. A quantitative approach was implemented to measure these concepts using a survey (N = 161). The data was gathered among adult YouTube users who consume entertainment content, whereafter the data was statistically analyzed using SPSS.

The results demonstrated that positive perceptions of the recommender system do not influence perceived entertainment content diversity, nor affect watchtime and neither influence the perceived amount of agency over the user's entertainment content consumption. However, more positive perceptions of the recommendation algorithm were found to lead to higher recommendation satisfaction. Also, content and technology gratifications were positively related to recommendation satisfaction, whilst social and process gratifications were not related.

There are three key takeaways that can be derived from this study. Firstly, the attitude that the user holds of the algorithm does not influence the amount of time users spend watching YouTube videos. So, regardless of whether a user has a negative perception of the recommendation algorithm, the user does not necessarily watch less YouTube which might be related to the privacy paradox theory. Secondly, an interesting outcome from this study that substantiates the algorithm appreciation theory is that positive perceptions regarding the recommender system lead to more recommendation satisfaction. Thirdly, users who experience convenience while using YouTube were also more satisfied with their entertainment recommendations. The second and third key takeaway justify the reasoning for digital media companies to optimize and refine their platform’s algorithm, user interface and user experience as much as possible since this leads to higher satisfaction and therefore possibly to higher platform usage.

The findings of this research contribute to the conclusions of previous studies and they might also be relevant to digital media companies, governmental entities and social actors.

KEYWORDS: *YouTube; User Perceptions; Artificial Intelligence; Recommendations; Uses and Gratifications*

Table of Contents

Chapter 1. Introduction

1.1 Research topic and research question

With over 2,291 million active users in 2021, YouTube is the social media platform with the second-highest number of users worldwide, right after Facebook (Statista, 2021). Besides this, as opposed to many other social media platforms that prove to be ephemeral (Arthurs et al., 2018), YouTube continues to expand, and it is the second most visited website after Google (Statista, 2022). The extremely popular social networking is specifically known for its video content sharing properties and its easy accessibility for both users and content creators (Balakrishnan & Griffiths, 2017). According to Balakrishnan and Griffiths (2017), the social networking site has multiple functions such as the provision of information, news, political messages and entertainment, which is one of the reasons why YouTube attracts so many users. The entertainment genre in particular has been the most popular genre since 2013 and videos in this category have an above average chance of getting in the top 3% of the most watched videos after uploading (Arthurs et al., 2018). Since entertaining videos are so prevalent and popular on YouTube, the research at hand will focus on the large share of videos that can be categorized in the entertainment category. In this study, the entertainment genre is defined as videos about gaming, sports, travel, pets and animals, music, shows, films, comedy, people and blogs, how-to and style and entertainment in general (Möller et al., 2019).

YouTube uses a proprietary algorithm that recommends new video content based on, among others, related videos, previous viewing behavior of the user and other users with similar viewing behavior (Covington et al., 2016). These recommended videos are clearly visible in the user interface (UI) on the YouTube homepage as well as near videos that are being watched, thus being a significant part of the user experience (UX) of the platform. The success of the recommender system seems evident since over 70% of the over 1 billion hours people spend watching videos on a daily basis is caused by the algorithm (Nicas, 2020). The algorithm is very effective in providing accurate recommendations and the vice president of engineering at YouTube even stated that video recommendations generate more views than both queries and channel subscriptions (Goodrow, 2021). Additionally, the main aim of YouTube as a business is to make users increase their watchtime, so users are exposed to as many advertisements as possible from which the platform earns its revenue (Arthurs et al., 2018).

The recommender system is a clear example of human-AI (artificial intelligence) interaction, where individuals are confronted with an intelligent medium that aims to understand and satisfy user needs whilst possibly affecting their content consumption behavior (Sundar, 2020). Concerning YouTube consumption behavior, a concept and buzzword that the platform has been connected to in scholarly works is the term ‘’filter bubble’’, which can be described as the possibly confined diversity within users' content recommendations, being caused by the algorithm (Roth et al., 2020). This means that users might only see recommendations of which the algorithm almost certainly knows that the user will continue watching, without providing new or oppositional information. Another aspect of YouTube consumption that has been studied is how much time users spend watching videos and whether some users may overconsume (Balakrishnan & Griffiths, 2017; Klobas et al., 2018). Possibly, the recommendation algorithm might be an influential factor to watchtime, because as mentioned before, the effectiveness of the recommender system is responsible for a large share of the overall YouTube consumption (Nicas, 2020).

Keeping in mind the important role that the algorithm plays in proposing videos to users, it might be interesting to explore whether YouTube entertainment content consumption is actually influenced by how users perceive the recommender system in the first place. Bucher (2017) has introduced the concept of the algorithmic imaginary to explore user perceptions of the Facebook algorithm and these perceptions may vary from being more positive to more negative. The concept of algorithmic imaginary can perhaps also be applied to the YouTube recommender system since the manner in which the algorithms filter content is similar for both platforms (Schafer et al., 2007). Academic literature discussing user perceptions of the YouTube recommender system is existent, yet in limited numbers. A previous study where YouTube usage was evident to be affected by perceptions of the recommender system has been executed by Bishop (2019). She encountered that beauty vloggers on the platform used their accumulated knowledge and experiences with the recommender system to enhance their channel performance. Hence, this is an example where perceptions of the YouTube algorithm were relevant to user behavior for a very specific group of content creators. However, a similar phenomenon where content consumption is affected by perceptions might also occur among YouTube users in general.

Moreover, varying levels of user satisfaction caused by YouTube consumption may also be interesting to connect to YouTube recommendations. This is because referring back to the affordance of AI in social media products to satisfy user needs (Sundar, 2020), this tendency can be associated with the uses and gratifications theory by Katz et al. (1973-1974).

This theory argues that users have active agency and awareness over their consumption habits and that all media gratify certain user needs through its platform, which as a consequence may affect content consumption habits and perceptions. For YouTube, these gratifications were distinguished as content, social, process and technology gratifications for which YouTube as a platform was encountered to satisfy different needs such as information provision or convenience (Bakar et al., 2014; Balakrishnan & Griffiths, 2017). However, none of these concepts have been studied with user perceptions of the recommender system in mind and that is where this research aims to make a difference. So, tying the discussed concepts in the context of YouTube entertainment consumption to user perceptions of the recommendation algorithm will be the goal of this study.

All by all, based on the aforementioned information this research aims to inquire about perceptions of the YouTube recommender system and how they affect entertainment consumption. Therefore, the following research question is proposed:

To what extent do perceptions of YouTube's recommendation algorithm shape user entertainment content consumption?

1.2 Societal and academic relevance

Regarding the societal relevance of this study, it must be acknowledged that millions of individuals use YouTube on a monthly basis (Statista, 2021). So, it might be valuable to have more knowledge of how the entertainment content consumption of these millions of people is shaped by such a vital, and as mentioned before, effective element of YouTube, which is the recommender system (Nicas, 2020). Also, existing biases among individuals may be reinforced by the filter bubble due to a lack of variation within the consumed content (Roth et al., 2020), which may lead to confined consumption behavior that limits the possible broadness of the entertainment scope of the individual. Furthermore, how users perceive the recommendation algorithm may be of interest to content creators on the platform. In December 2021 YouTube decided to remove the dislike button, to reduce hateful actions against small content creators (Suciu, 2021). However, according to Suciu (2021) small content creators actually feel disadvantaged by this change and they would prefer the old system. This is one example of how the systems created by YouTube and the affordances of the platform affect and interact with society and it might be interesting to gain a more in-depth understanding of the user effects of the recommendation algorithm. Besides content creators, digital media companies may also benefit from having further insights on how perceptions of comparable AI implementations affect content consumption. Companies such as Spotify, Netflix and Instagram that work with algorithms similar to YouTube might be interested in the outcomes of this study to gain a deeper understanding of the behavior of their users. Lastly, the results of this study certainly might be interesting to YouTube itself for the same reason.

Concerning the academic relevance of this research topic, several points might be interesting and essential to inquire about. Relating back to entertainment content, a study discovered that stronger motivation to use the platform for entertainment leads to higher compulsive usage (Klobas et al., 2018). Thus, perhaps a user who continues to receive perfectly tailored recommendations might be more inclined to compulsively consume entertainment videos, creating a spiral-effect since the quality of recommendations were found to only improve overtime (Bryant, 2020). However, in academic research the role of perceptions of the recommendation algorithm has not been taken into account in this context yet. Moreover, in existing literature users were found to have a lack of understanding and awareness of the workings of the recommendation system (Alvarado et al., 2020). Alvarado et al. (2020) mention how middle-aged users were somewhat aware of the existence of the algorithm, yet they only had superficial understanding of the technology. Hence it might be interesting to further study the aforementioned concept of algorithmic imaginary of YouTube users in the general population. Moreover, the uses and gratifications theory in the context of AI implementation in social media is an understudied subject as well. Thus, the academic relevance of the proposed research concerns the gap in existing literature on this subject.

1.3 Structure

In the upcoming chapters, the reasoning behind the answer to the proposed research question is developed by exploring the existent body of literature, executing a quantitative research method and presenting and discussing the encountered outcomes. More precisely, chapter two presents an overview of previous research on the YouTube recommender system and it will further discuss the studied concepts of the algorithmic imaginary, watchtime, perceived content diversity, the uses and gratifications theory and recommendation satisfaction. The third chapter discusses the research method as being a survey and how the measured concepts were operationalized, whilst checking for reliability and validity. The fourth chapter provides the results that were found after conducting statistical analyses. Chapter five is devoted to the discussion of the outcomes in relation to existing academic works as well as societal implications, limitations and strengths of the study and suggestions for future research. Finally, chapter six concludes this research by stating the concluding remarks with three key takeaways from the study.

Chapter 2. Theoretical framework

In the following chapter, theoretical concepts will be discussed to create the framework surrounding user perceptions regarding the YouTube recommendation system and how this shapes entertainment content consumption. First, background information on YouTube and its recommender system is provided. Second, algorithmic imaginary and other related concepts are analyzed. Hereafter, content diversity, watchtime, recommendation satisfaction, and perceived agency over consumption are discussed with respect to perceptions regarding the recommendation algorithm. Third, building from the uses and gratifications theory, the role of the four user gratifications, content, social, process, and technology, is addressed in relation to the user's recommendation satisfaction. To conclude, a conceptual framework including a hypothetical model is presented where the relationships between the presented concepts are visualized.

2.1 YouTube and its recommendation system

The well-known social networking site YouTube was launched in 2005 and has had millions of users worldwide ever since (Abbas et al., 2017). The social media platform, which is now owned by Google (Arthurs et al., 2018), has been the second most popular website since 2019, which clearly shows the impact that YouTube has on society (Roth et al., 2020). YouTube provides users with video content and largely consists of user-generated content (UGC), as well as the possibility for interaction with creators and other users by liking, disliking, commenting, and sharing on the platform (Khan, 2017). Considering user motives for YouTube usage, 79% of the teenagers on the platform were found to access YouTube to satisfy their need for entertainment (Chau, 2010) and in a more general study, users who watched YouTube videos in a passive manner were more likely to do so for relaxing entertainment reasons (Khan, 2017). This also strengthens the finding that entertainment has been the most popular category on YouTube since 2013 (Arthurs et al., 2018). Cunningham and Craig (2017) discuss the framework around social media entertainment (SME) on YouTube that is based on the content of former amateur creators who have professionalized themselves on the platform by making videos about vlogging, gameplay, and style tutorials. In addition to this, Möller et al. (2019) studied user responses to videos in the entertainment and political genre. They however argued for a broader entertainment scope existing of videos categorized in the topics of entertainment, music, films, how-to and style, gaming, shows, comedy, pets and animals, people and blogs, travel, and sports (Möller et al., 2019). Thus, entertaining content can be considered a key component of the YouTube video library.

Whilst browsing through the platform, YouTube users have a seemingly infinite amount of content at their disposal (Abbas et al., 2017). In the early days of YouTube, the homepage automatically consisted of the most popular videos at that moment such as new music videos or sports fragments (Goodrow, 2021). It is however imaginable that these videos were not in line with the actual interests of most users, so most views came from shared links or specific searches. Consequently, to help users sift through the enormous video collection and find what really interests them, the recommender system was implemented in 2008 (Goodrow, 2021). This system is based on machine learning which can be described as the ability of computational technologies to find underlying patterns in data and infer underlying rules accordingly (Sundar, 2020). These rules can then be applied autonomously by a system using an algorithm that is taught to reach a specific goal such as offering a recommendation, which is called AI. Internal researchers at Google in 2010 (Davidson et al.) state that "the goal of the system is to provide personalized recommendations that help users find high-quality videos relevant to their interests." (p. 293). Davidson et al. (2010) also mention that 60% of the views from the homepage were already caused by recommended videos, proving the system to be successful. In 2021, recommendations were found to generate more views than channel subscriptions and queries together (Goodrow, 2021). Over the past years, the recommendation UI, as well as the recommendation system, have been developed further. For example, in 2012 rather than only noting whether the user has clicked on a video, the algorithm started taking watchtime into account to measure whether the user actually pursued watching. Also, in 2016 users were given the option to rate videos after watching them on a 5-star scale to determine the extent to which users actually enjoyed the video (Goodrow, 2021). After these algorithmic developments amongst others, researchers at Google even went as far as calling it the most sophisticated industrial recommendation system in existence (Covington et al., 2016).

As stated before, the recommendation system is part of the algorithm engineered for YouTube and the computational principle is based on deep learning neural networks (Abbas et al., 2017; Covington et al., 2016). More precisely, YouTube applies a hybrid recommendation approach that considers both content-based recommendations as well as collaborative filtering (Covington et al., 2016). Regarding content-based recommendations, these recommendations are based on videos that the user has watched in the past (Abbas et al., 2017). Here, the similarity of the content that is being watched is of importance and this is mostly evaluated using the thumbnail, title, or description of the video. As for collaborative filtering, Schafer et al. (2007) describe this as using the profiles of users with similar user behaviors to predict the likelihood that a certain user will be interested in a certain recommendation. So, items, being videos in this case, are filtered using the ratings of other individuals. Hence, both approaches are implemented into the YouTube recommendation algorithm, making the system hybrid (Abbas et al., 2017). The algorithm improves through machine learning, so every time successful interactions occur when a recommended video is actually watched, the algorithm learns that there is a positive relationship between the watched video and the recommended video (Bryant, 2020). The precise engineering details of the algorithm are somewhat of a mystery, however, since Google like most large online platforms does not disclose its computational formula for economic reasons (Arthurs et al., 2018). This is called a "black box" algorithm that leaves outsiders with many questions about the exact formulation of the recommender system (Bryant, 2020).

As mentioned earlier, the need for entertainment is the primary motive for watching YouTube videos (Arthurs et al., 2018; Chau, 2010; Khan, 2017). Existing literature regarding the YouTube recommendation system often examines the issue in a political and often extremist radicalized context (e.g., Bryant, 2020; Haroon et al., 2022; Munger & Philips, 2022). However, the focus of interest of this research entails the entertainment genre in the broadest, most applicable sense as discussed by Möller et al. (2019) and aims to create a better understanding of how the recommendation system shapes entertainment content consumption. The remainder of this chapter will address several concepts that may play a role in this phenomenon.

2.2 The algorithmic imaginary

Rather than solely looking at algorithms from a computational perspective, it is also of importance to regard them from the perspective of social embeddedness, as is currently more often done by social scientists (Schellewald, 2022). According to Sundar (2020), media studies are shifting from a Human-Computer interaction approach, where humans treat computers as autonomous social actors, to a Human-AI interaction approach, where users view the medium as intelligent and capable of modifying content in revolutionary ways. The occurrence of the Human AI-interaction approach is prevalent in many digital technologies nowadays, for instance, social networking sites such as Facebook or smart speakers like Amazon's Alexa (Sundar, 2020).

When users interact with AI and its algorithm, they create certain perceptions about said algorithm. Taina Bucher (2017) most notoriously describes this as an individual's algorithmic imaginary, which she defines as "the way of thinking about what algorithms are, what they should be, how they function and what these imaginations in turn make possible" (p. 40). The author discusses this concept in the context of Facebook's algorithm and how users experience and make sense of the algorithm, as well as the extent to which users are aware of its existence. Bucher (2017) cataloged while taking a qualitative approach, six reactions to the impact of the algorithm that Facebook users experienced while using the platform. For instance, what she calls the "Whoa moments" is a user reaction when encountering that the user has been noticed by the algorithm, by for example seeing an ad of the exact brand of coffee the user is drinking at that time.

The concept of algorithmic imaginary has been implemented into the works of multiple media scholars (e.g., Bishop, 2019; Schellewald, 2022). Schellewald (2022) used the concept of algorithmic imaginary to do ethnographic research on the algorithmic experience of TikTok users while scrolling through their content feed. In addition to this, Bishop (2019) described the occurrence of "algorithmic gossip" that occurs when content creators gain and share knowledge of the processes behind algorithms and how to use this knowledge to their benefit in the context of YouTube beauty vloggers. For this theory, Bishop drew from both algorithmic imaginary and the so-called "folk-theories". The latter is another concept in the realm of algorithm awareness that maps the variety of existing folk theories created by users about a specific recommender system (Bishop, 2019; DeVito et al., 2017). Moreover, Alvarado and Wearn (2018), opted to focus more specifically on how users experience systems and platform interfaces in which the algorithm is incorporated. They conceptualize this as Algorithmic Experience (AX), creating a framework to make human interactions with algorithms explicit. It should be noted that the AX approach is centered around the social media platform's affordances and how the algorithm is integrated into its interface, focusing less on the subjective user perceptions of the algorithm. Considering beliefs that the users hold of the YouTube algorithm more specifically, Alvarado et al. (2020) performed interviews with middle-aged users to study this. They found that users accredited their video recommendations to four actors: the algorithm, the organization itself (YouTube), the current users, and other users. Furthermore, these middle-aged users were found to be aware of the algorithm whilst still having a limited understanding of its mechanisms. In short, different forms of algorithmic awareness among social media users have been studied in various contexts already, so it is reasonable to inquire further and more specifically about user perceptions of the YouTube recommender system.

As mentioned before, the algorithmic imaginary is described as the manner in which all social media users view algorithms and how they shape consequential user behavior (Bucher, 2017), which relates directly to the posed research question of the study at hand and is therefore the most relevant conceptual framework. According to Sundar (2020), the medium that shapes the media user was previously a constant in its characteristics. Contemporary media, however, such as social media platforms, are mediated by AI technologies of which the workings are constantly modified through deep learning and system changes, creating new affordances. Thus, media researchers can no longer treat the technology of a certain medium as a constant that shapes users, and they should continuously explore new phenomena in the field of Human AI-interaction (Sundar, 2020). Moreover, Bucher (2017) states that understanding how algorithms make people feel is essential in understanding the social power that they hold. Currently, there is a gap in the literature concerning the algorithmic imaginary in the context of YouTube as well as this concept in relation to other factors shaping entertainment content consumption on YouTube. Therefore, the following subsections will explore user perceptions regarding the YouTube recommender system in relation to content diversity, watchtime, recommendation satisfaction, and perceived agency over consumption, all in the context of entertainment content.

2.2.1 Effect on perceived recommendation content diversity

The goal of the recommendation system is to provide personalized recommendations that make users continue watching (Bryant, 2020). Researchers at Google claim that when video consumption increases, more specialized suggestions can be made according to the user's interests (Baluja et al., 2008). This idea relates to the concept of the filter bubble that was most famously coined by Eli Pariser (2011). This concept is described as the phenomenon where recommendation algorithms cause users to solely be exposed to information reinforcing their own viewpoints, without receiving contradicting information. Through the concept of the filter bubble, Pariser (2011) explains that the lack of content diversity might be accompanied by a lack of opposing views whilst consuming content which leaves individuals in a rabbit hole of restricted information provision that might be problematic. On YouTube, filter bubbles were also found to be apparent (Roth et al., 2020). Here, recommendations were found to be susceptible to confinement dynamics, and the most confined bubbles were encountered among the most popular videos with the highest number of views. Bryant (2020) studied the filter bubble in the context of extreme right-wing ideology videos on YouTube and states that the algorithm has a strong bias towards right-wing political videos and leads users to a rabbit hole of radical information without providing opposing views. Thus, the aim of recommendations may be to incite users to continue watching, an additional consequence might be that users may be incentivized to radicalize due to the filter bubble (Bryant, 2020). Corporate workers at YouTube are allegedly aware of this issue and act on it by demoting what they call borderline content, being extremist information or conspiracy theories among others, within the recommendation system (Goodrow, 2021). According to analyses from within YouTube, this has led to a 70% decrease in the consumption of recommended borderline videos in the US. Thus, YouTube attempts to combat the influence of filter bubbles, yet the issue persists to be perceived (Bryant, 2020).

The concept of the filter bubble has predominantly been addressed in the context of political videos on YouTube (Bryant, 2020). Yet, due to the user-specific engineering behind neural networks, it is reasonable to assume that a similar restrictive phenomenon occurs in the entertainment genre. This assumption is made since researchers at the music platform Spotify, which uses comparable forms of machine learning for their algorithms, found that content diversity reduced when users did more algorithmically-driven listening through their tailored recommendations (Anderson et al., 2020). Additionally, it was observed that users who diversify their content consumption, do so by shifting away from algorithmic consumption and moving towards organic consumption by finding content through queries for example.

Studies on the diversity of entertainment content recommendations on YouTube are lacking and it may be interesting to study this in relationship with perceptions regarding the recommendation algorithm. It is likely to assume that users who view the YouTube recommendation system more positively, will be less likely to be aware of the filter bubble so they might be more inclined to fall into the rabbit hole of the system and therefore have less diverse entertaining recommendations. Building onto the aforementioned arguments, the first hypothesis is stated:

H1: Positive perceptions regarding the recommendation algorithm lead to less diverse entertainment content recommendations.

2.2.2 Effect on watchtime

Aside from the diversity of the consumed entertainment content, the quantity of consumption might also be intriguing to take into consideration. Studies on the subject of social media consumption time mostly focus on Facebook (Kuss & Griffiths, 2017). Studies on YouTube consumption time are underrepresented and the few existing ones target the extremes of content consumption: addiction and compulsive usage (Balakrishnan & Griffiths, 2017; Klobas et al., 2018). Due to the many affordances of YouTube, such as entertainment and information provision, chances of becoming addicted to the platform are equal, if not higher, than becoming addicted to mainstream television (Balakrishnan & Griffiths, 2017). Furthermore, users who create content were also found to have a higher score on the addiction construct than users who only watch videos on the platform. Klobas et al. (2018) even claim that watching entertainment video content on YouTube had a significant effect on addiction to the platform. This study, which was conducted among Malaysian students, found that almost 20% of the students report a lack of control over their consumption resulting in compulsive usage and males were also more likely to use the platform excessively (Klobas et al., 2018).

The evaluation of the success of the YouTube recommendation system partly stems from the user's watchtime, where a higher watchtime is deemed more successful (Covington et al., 2016). So, the algorithm will attempt to make the user watch as many videos as possible. O'Donovan et al. (2019) even go as far as calling the algorithm an ‘’engagement monster’’ that will go to the greatest lengths to keep users watching a few more seconds. In the existing literature, the relationship between a user's algorithmic imaginary and their amount of entertainment consumption has not been explored yet, also not in the context of YouTube. It might be likely that individuals who view the recommendation system more positively are also likely to consume more content since they trust the quality of the recommendations that are given to them. So, because of the aforementioned reasons, the second hypothesis is proposed:

H2: Positive perceptions regarding the recommendation algorithm lead to more time spent consuming entertainment content.

2.2.3 Effect on recommendation satisfaction

It has been discussed how users may experience algorithms on social media and how they are aware of their existence through algorithmic imaginary. Next, user satisfaction with regards to their recommended videos in relation to how they perceive the functioning of the recommendation algorithm will be analyzed. This connects to the concept of algorithm aversion, which is defined as the human distrust of algorithmic output (Dietvorst et al., 2015). Dietvorst et al. (2015) explain that in their experiment, humans were more likely to avoid trusting the judgment of AI after it makes a mistake than to avoid trusting a human's judgment after making the exact same error. This is because individuals have higher expectations of AI and are therefore more reluctant to confide in AI again after seeing it err, even after seeing it outperform an individual. More specifically, humans may show more algorithm aversion on some tasks than on others as was discovered by Castelo et al. (2019). The authors found that consumers are less likely to rely on tasks that are subjective in nature, such as suggesting items while online shopping, as opposed to objective tasks.

On the other hand, besides algorithm aversion, algorithm appreciation is also a phenomenon, albeit less prevalent in the literature (Logg et al., 2019). According to this notion, humans prefer the judgment of AI over human judgment. Logg et al. (2019) discovered that algorithm appreciation occurred when romantic attraction and the popularity of music were forecasted. In relation to this, a study from 2019 (Banker & Khetani, 2019) found that some consumers often depend too much on algorithm-generated recommendations when shopping online, even when these recommendations were inferior. This is conceptualized as algorithm overdependence and may result in a negative impact on one's well-being and reinforce systematic bias (Banker & Khetani, 2019). On the contrary, professional experts who make predictions on a regular basis were found to rely on AI less when making decisions, which eventually led to poorer personal predictions (Logg et al., 2019). These contradicting findings show that hubris both regarding AI and human advice should be received critically and that levels regarding recommendation system satisfaction may vary among users.

Specific information on how satisfied users are with their entertainment content recommendations on social media is not available. Building onto the aforementioned information, it is reasonable to assume that individuals who hold a more positive attitude regarding the recommendation system will also be more satisfied with the recommendations they receive (Banker & Khetani, 2019). This is likely because users will then pass their positive experience with the algorithm onto its products, the recommendations, and trust the algorithmic process more. Hence, the third hypothesis is as follows:

H3: Positive perceptions regarding the recommendation algorithm lead to higher user satisfaction concerning their entertainment content recommendations.

2.2.4 Effect on perceived user agency

Many scholarly works that are focused on audience perceptions of media products have related this to the uses and gratifications (U&G) theory. This theory was coined in the 1940s, yet is most notoriously attributed to Katz et al. (1973-1974). The U&G theory claims that mass media users utilize media to gratify certain wants and needs, hence the name. These needs are characterized by two principles that are assumed regarding the media user. Firstly, media users are seen as active agents who make conscious and controlled choices when consuming media. This principle is in sharp contrast with previous media theories such as the hypodermic needle model that assumes that users passively consume media content and receive messages without any interaction between the individual and medium involved (Nwabueze & Okonkwo, 2018). The second principle states that users are completely aware of the reasoning behind their content consumption choices since they want to gratify their wants and needs (Katz et al., 1973-1974). Other well-known media theories are the social cognitive theory and the cultivation theory. These are however deemed to be less fitting in the scope of this research since social cognitive theory aims for a more personal application of the U&G theory (Bandura, 2005) and cultivation theory is more focused on how media shapes one's worldview and the consequential psychological effects on a more longitudinal scale (Potter, 2014). One of the goals of this study is to create an understanding of the current day effect of user perceptions on the media consumption of a larger audience, making the U&G theory most appropriate (Katz et al., 1973-1974).

Before diving into the gratifications of the U&G theory, first the aspect of user agency will be explored. In contemporary works, the U&G theory has been related to the agency that algorithms exert in contrast with human agency by academics Sundar and Marathe (2010). The authors discuss that in their study users who were more tech-savvy and thus experienced in using communication technologies preferred having a self-tailored news feed over a personalized one. This is because these individuals tend to be more skeptical of privacy concerns and therefore prefer customizing their news consumption. Users who were less experienced with new media on the other hand preferred news feeds that were personalized by the algorithm. This shows that the preferred amount of agency in the diversity of content consumption may vary (Sundar & Marathe, 2010).

Relating back to the work by Katz et al. (1973-1974), the U&G theory stems from a time pre-current of social media platforms and machine learning implementation altogether. Nonetheless, it may be interesting to explore how the principles of the theory hold with regard to the context of the research at hand. Several other studies in the new media field also implemented the U&G theory when exploring the social implications of YouTube (e.g., Bakar et al., 2014; Balakrishnan & Griffiths, 2017; Khan, 2017). Yet, there is still a gap in the literature concerning perceived agency over user's entertainment content consumption in relation to user perceptions regarding recommendation systems, which is where this study aims to make a contribution. It is likely to assume that some users may experience a sense of loss of control over choices of media consumption caused by the YouTube recommender algorithm. This may occur when users dislike the "pushy" tendencies from the recommendation system and a user may want to avoid being influenced against their will. Therefore, the fourth hypothesis is posed:

H4: Negative perceptions regarding the recommendation algorithm lead to a lower perceived amount of agency over the user's entertainment content consumption.

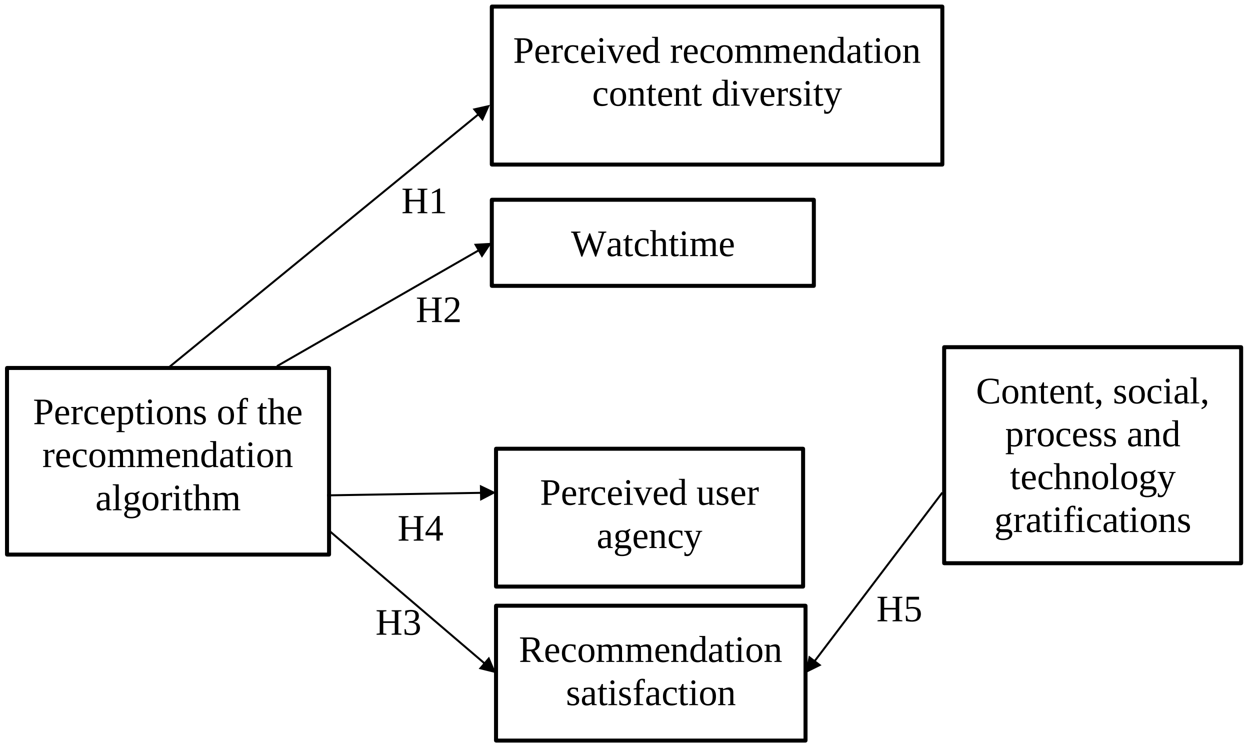
2.3 YouTube user gratifications

As mentioned in the previous subsection, users have certain needs that can be gratified using mass media platforms (Katz et al., 1973-1974). The research by Katz et al. (1973-1974) which mostly focuses on television, radio and print media as the covered mediums, discussed information provision, entertainment, using media to pass time and enhancing social interactions as the main user gratifications. In studies concerning YouTube consumption, four types of user gratifications are distinguished: content, social, progress, and technology (Bakar et al., 2014; Balakrishnan & Griffiths, 2017). More precisely, content gratification entails the need for information-seeking by users, which is the most important need for most users on any media platform (Bakar et al., 2014). If a user perceives the quality of the content that is provided in a YouTube video to be adequate, they will have a high content gratification score. Secondly, social gratifications are addressed when social needs are met and users can engage in interactions with other users by commenting, liking, and sharing videos (Balakrishnan & Griffiths, 2017). When users experience a high sense of connection with others, they will score high on the social gratification scale. Thirdly, process gratification addresses the effective utilization of the medium. So, users are gratified by being involved in the practical process of the medium, rather than by the content of the information they encounter (Bakar et al., 2014). If a user experiences a feeling of entertainment or when they pass time by watching YouTube videos, they score high on the process gratification scale. Fourthly, technology gratification concerns the medium appeal and convenience with which users can utilize the social media platform (Balakrishnan & Griffiths, 2017). Hence, users who are satisfied with the online environment provided by the YouTube platform will score high on the technology gratification scale. So, the four distinguished YouTube gratifications are somewhat similar to the ones postulated by Katz et al. (1973-1974).

In existing academic literature, the possible influence of the four types of gratifications on user satisfaction regarding their entertainment content recommendations has not been analyzed yet. User needs that are gratified will probably lead to higher user satisfaction when interacting with the YouTube platform and encountering recommended videos. For instance, a user who is satisfied with the informational nature of the videos they consume (content gratifications), may also be satisfied with their recommendations, since these are likely to provide this adequate information. Thus, it is likely to consider that a high content, social, process, and/or technology gratification score will affect user recommendation satisfaction. Therefore, the fifth and final hypothesis is posed:

H5: The user's (a) content, (b) social, (c) process, and/or (d) technology gratifications score positively affect the user satisfaction score regarding their entertainment content recommendations.

2.4 Conceptual framework



In short, this chapter addressed how user perceptions of the recommender system are discussed in the existing body of literature and how these perceptions might affect perceived content diversity, watchtime, perceived user agency and recommendation satisfaction. Moreover, the influence of user gratifications on recommendation satisfaction is mentioned. Also, five hypotheses regarding these subjects were proposed. The following chapter will discuss the research method and operationalization of the discussed concepts.

Chapter 3. Method

In the following chapter the methodological approach of this study will be addressed. The aforementioned hypotheses will be investigated and reasons for executing a quantitative study will be described. Then, the research design will be proposed as being an online survey to collect data. Hereafter, the descriptive statistics are discussed to illustrate the studied sample. Next, the mode of analysis is further clarified by discussing the operationalization of the variables. Finally, the reliability of the variables is addressed by conducting factor and reliability analyses which is accompanied by a discussion of the validity of the research design.

3.1 Methodological approach

The purpose of this study is to inquire about the perceptions that users have of the YouTube recommender system and how this affects watchtime, perceived content diversity, recommendation satisfaction and perceived user agency. Additionally, the relationship between content, social, process and technology gratifications and recommendation satisfaction is studied. Hence, this research seeks to discover how one variable influences the other in an empirical matter through hypothesis-testing. For this reason, a quantitative research design was applied, since a correlation between independent and dependent variables can be found through this approach (Salkind, 2010).

To reach the goal of this study, the instrument that was applied was an online survey. A survey is an appropriate means to collect data for an exploratory study according to Babbie (2011) and since user perceptions of the YouTube recommendations system are still relatively untrodden ground, this method seems fitting. Also, this method allows for possible representativeness across the general population (Salkind, 2010), which would be interesting considering the significant amount of YouTube usage globally (Roth et al., 2020). A qualitative method would not be suitable for this study, since qualitative research is more generally executed to find patterns in smaller data sets to create an in-depth analysis of a specific phenomenon (Braun & Clarke, 2006). Whereas this research aims to study the larger population and search for generalizable findings. Thus, in order to answer the research question an online survey was deemed to be the most appropriate tool.

3.2 Research design

Using the research software Qualtrics, an online survey was designed to execute this study. The survey was distributed by the researcher through online platforms which is fitting because the participants needed to be active online (using YouTube) to belong to the target group.

In academic literature, the algorithmic imaginary as proposed by Bucher (2017) is mostly studied through qualitative ethnographic studies to research user experiences from close by (e.g., Bishop, 2019; Schellewald, 2022). Alvarado et al. (2020) on the other hand opted to perform interviews to study user beliefs about the YouTube algorithm, which is also a qualitative and more in-depth method. However, the study at hand aims to analyze user perceptions on a larger more generalizable scale since YouTube usage is so ubiquitous and therefore an online survey seemed the most appropriate. Also, online surveys have already been used by other researchers when they studied YouTube and its user motives (Khan, 2017) as well as addiction (Balakrishnan & Griffiths, 2017) and compulsive usage of the platform (Klobas et al., 2018). This demonstrates that using an online survey is an effective method to inquire about YouTube consumption and find relevant results. Hence, this research method is quite novel in this realm, yet it may be a good contribution to the existing body of literature.

Additionally, since the U&G theory is part of this study, an online survey is again an appropriate method because the U&G theory focuses on the behavior of individuals in a short time span (Katz et al., 1973-1974). Using survey questions, this exact indication of user behavior can be measured in that certain point of time. This way, the current day effect of the YouTube recommender system can be analyzed, since only multiple surveys would allow for measuring long term effects which is outside the scope of this research. Next, the survey structure and the data collection will be addressed.

3.2.1 Survey structure

The survey, which is presented in appendix A contains six parts. In the first part the participant is introduced to the subject of the study and it is explained how participating is on a voluntary basis. Also, the confidentiality of the gathered data is addressed, whereafter participants have to give their consent to start taking the survey. Hereafter, a filtering question was posed which asked whether participants have used YouTube in the last month to watch entertaining videos. If the participant answered "no" the survey was terminated. Then, participants were asked to enter their age so people under the age of 18 could be led to the end of the survey. Next, users indicated how much time they spend watching YouTube videos per week. Hereafter, using matrix tables participants rated their opinions on statements on the concepts perceived user agency, perceived recommendation content diversity and recommendation satisfaction. Following, participants indicated their perceptions of the recommendation system and how the platform YouTube gratifies their needs. At last, participants were asked to enter their demographic information being gender, level of education and nationality as well as additional information on whether they are content creators on YouTube and which devices they use to watch YouTube videos.

Participants were asked to answer questions about the dependent variables first and the independent variables afterward. This is because measuring the dependent variable first prevents the user responses from being conditioned by the independent variable, so the effect between the two variables is measured more accurately. The demographic and additional questions were posed at the end to combat survey fatigue. The constructs that were measured to answer the hypotheses were all rated on a five-point Likert scale to create internal consistency as well as for the convenience of the participants. Participants were found to be more likely to take surveys from their smartphone (Morgan, n.d.), so a five-point scale would be more user-friendly in the interface of the survey. Moreover, the statements from the questions answered with a Likert scale were randomized to increase validity and eliminate order bias among participants. In §3.4 the variables that were measured in the survey are elaborated upon further.

3.2.2 Data collection

In the initial stages of the creation of the survey, the researcher created her own scale for the variable *perceived recommendation content diversity.* A pre-test with a sample of 17 participants was conducted to test whether the scale was reliable. The scale had an unreliable Chronbach's alpha which resulted in the researcher using an existing scale by Willemsen et al. (2016) in the final version of the survey instead. In addition to this, prior to distributing the complete survey another pre-test was conducted to enhance the chances of a correct interpretation of the questions and scales and omit errors in the overall flow, as recommended by Babbie (2011). The pre-test sample consisted of two male and three female participants and after each pre-test the researcher and participant discussed the comprehensibility of the overall survey and points that needed improvement. One participant noticed that the survey initially asked participants for their watchtime recalculated in minutes. This participant however pointed out that she, and possibly many more YouTube users, has a watchtime of about 11-13 hours per week and that it might be too much effort for participants to recalculate their watchtime. Therefore, this question was adjusted so that participants could simply indicate their weekly watchtime as they wished. Furthermore, some adjustments in wording of the statements were done to increase clarity.

The desired sample size of this research was 150-250 participants and Pallant (2016) recommends a sample size higher than 150 when conducting factor analyses. Hence, the target sample size of this research is > 150. Regarding the sampling procedure of this study, the nonprobability techniques snowball and convenience sampling were used (Babbie, 2011). These sampling methods are known for their practicality and convenience since participants are sampled based on their availability. This may however also lead to a similar and biased sample that is more difficult to control (Babbie, 2011). For the purpose of this study however, that is not political in nature and focused on a digital platform that is so widely used in the population, chances of bias are lower and snowball and convenience sampling are deemed appropriate. However, there is a possibility that these sampling strategies may condition the generalizability of the encountered results.

Participants needed at least some understanding of the digital world to fulfill the criteria of participating in the survey as online platforms were used as the primary means for distribution. Thus, accompanied by a short introduction of the content and format of the survey, the survey was distributed through multiple online channels being WhatsApp, Facebook, Instagram, LinkedIn and Reddit. On WhatsApp the survey was shared with connections of connections of the researcher. Direct friends and family of the researcher were clearly instructed not to take the survey to prevent biased results. Also, on social media sharing of the survey was encouraged to reach an as large and diverse population as possible. Moreover, the platform SurveySwap was used as well as three public Facebook groups where surveys were exchanged, mostly among students. Using these platforms likely increased the diversity of the sample, since users in these Facebook groups were from all over the world, studying at different universities. On the other hand, utilizing these platforms may have also led to more bias in the sample since mostly higher-educated students are active in these Facebook groups and on SurveySwap which may lead to an overrepresentation of a certain age category as well as individuals with a higher-educated background.

3.3 Descriptive statistics

During the period of May 3, 2022, until May 11, 2022, in total 233 individuals opened the survey. However, 71 people did not complete the entire survey, and were therefore excluded from the final dataset. These participants stopped at varying points during the survey, yet most of them stopped at the third matrix question, which might mean that the survey took too long for them. A threshold for the duration time of the survey was calculated to determine whether participants spent enough time properly filling in the questionnaire. This was examined using a standard as discussed by Soland et al. (2021) by taking the median duration time, which was 332 seconds and multiplying this by 0.40 which results in 132.8 seconds. Then, one more participant was excluded from the dataset since they took 120 seconds to complete the survey, which probably means they did not complete the survey seriously since they scored under the threshold of 132.8 seconds. Thus, after data cleaning the final dataset consisted of 161 participants. For the watchtime question however one participant could not indicate their watchtime, so for that construct $N = 160$.

3.3.1 Descriptive statistics: Respondents

The next subsection will specify the descriptive demographic statistics from the respondents. The variable *age* was stated as an open text question and. The age of the participants ranged from 18 to 67, with $M = 26.07$, $SD = 8.62$. Concerning *gender*, the sample consisted for 64.6% ($N = 104$) of females, 31.7% ($N = 51$) of males, 3.1% of non-binary individuals and 0.6% ($N = 1$) of the responses was missing. Regarding highest completed *educational level*, two groups represented most of the respondents, being university bachelor's degree with 39.8% ($N = 64$) and university of applied sciences or similar with 19.3% ($N = 31$). The rest of the sample indicated that they finished a graduate or professional degree (14.9%, $N = 24$), post-secondary vocational education or similar (14.3%, $N = 23$) and middle school (9.9%, $N = 16$). For primary school, prefer not to say and the missing values respectively one individual indicated this which counts for 1.8% in total. The descriptive statistics for gender and educational level are presented in table B1 in appendix B.

Considering the participants' *nationality*, the majority of the sample was Dutch (52.8%, $N = 84$), followed by the UK (7.5%, $N = 12$) and the US (6.9%, $N = 11$). The full list of nationalities is presented in appendix table B2.

3.3.2 Descriptive statistics: YouTube usage

One of the requirements for partaking in the survey was that the participant should have been active on YouTube during the last month and watch entertainment content. So, all participants are active YouTube users. The average weekly *watchtime* for entertainment content was 302.16 minutes (*SD* = 407.38) with a minimum of 1 minute and a maximum of 3136 minutes. It was also questioned how aware participants felt of the tracking behavior of the algorithm with 1 meaning very unaware and 5 being very aware. For this item the average score was 4.35 (*SD* = 0.66). Furthermore, 16 participants (9.9%) indicated to be content creators on YouTube and the rest did not (90.1%, *N* = 145,). Finally, 150 (93.2%) of the participants watch YouTube on their smartphone, 126 participants (78.3%) use their laptop/desktop, 63 participants (39.1%) use their TV screen, 28 participants (17.4%) use a tablet and in the text entry for "other" one (0.6%) participant indicated to use a game console and one other person (0.6%) used a smart watch and VR set. Appendix table B3 demonstrates the descriptive statistics of the rate of content creators among the sample and device usage.

3.4 Operationalization

Gender. Participants were asked to indicate their gender affiliation (1 = Male, 2 = Female, 3 = Non-binary, 4 = Prefer not to say).

Age. Participants were asked to indicate their age in numbers (in full years).

Educational level. Participants indicated their highest completed educational level by choosing a single option (1 = Primary school, 2 = Middle school, 3 = Post-secondary vocational education or similar, 4 = University of applied sciences or similar, 5 = University Bachelor's degree, 6 = Graduate or professional degree, 7 = Prefer not to say).

Nationality. Participants indicated their nationality by opting from a list of 193 countries.

Devices. Participants were asked to indicate which device(s) they use to watch YouTube videos with "Smartphone", "Laptop/desktop", "Tablet", "TV screen" or "Other" with a text entry possibility as the answer options. The selection of multiple answer options was possible.

Content creator. Participants indicated whether they also create content on YouTube by answering "Yes" or "No" to the question "Do you also create content on YouTube yourself on a regular basis? So, do you upload videos on YouTube?".

Perceptions of the recommendation algorithm. To measure this concept existing scales were lacking, however the six reactions to awareness of the Facebook algorithm that

Bucher (2017) categorized were used by the researcher as a guideline to create a new scale. This scale contains five items (e.g., "I get frustrated or annoyed when the recommendation algorithm shows me videos that I am not interested in", "I think it's convenient that the algorithm understands my interests and suggests the right videos to me".). The items were scored using a Likert scale ranking ranging from 1 *(strongly disagree)* to 5 *(strongly agree)* to create internal consistency between the questions. The overall level of perceptions of the recommendation algorithm was determined by calculating the average score across the items (M = 3.28, SD = 0.64; Cronbach's α = .70).

Algorithmic awareness. Again, based on the categorizations by Bucher (2017), participants were presented with the following statement: "I am aware that the algorithm tracks my video watching behavior". This item was scored using a Likert scale ranking ranging from 1 *(strongly disagree)* to 5 *(strongly agree)* and the overall degree of algorithmic awareness was measured by calculating the average score (M = 4.35, SD = 0.66).

Perceived recommendation content diversity. To measure the perceived recommendation content diversity a scale implemented by Willemsen et al. (2016) was used. Four out of five items from this scale were adapted and adjusted in wording to fit the research scope. This scale has also been used in other studies by Ferwerda et al. (2017) and Lee and Lee (2022) to study diversity in recommendation systems, which makes this scale more valid and appropriate to implement. Examples of the items are "My list of recommended videos is varied and diverse" and "All of my recommended videos are similar to each other". The statements were rated using a five-point Likert scale varying from 1 *(strongly disagree)* to 5 (*strongly agree).* The overall level of perceived recommendation content diversity was determined by calculating the mean score across the items (M = 2.99, SD = 0.75; Cronbach's α = .77).

Watchtime. To measure watchtime, participants were asked the following questions: "How much time do you spend watching entertainment videos per week?". It was recommended to participants to check their watchtime in the YouTube app for a more accurate estimate. Participants were however urged to be cautious with using the watchtime statistic from the YouTube app since this shows the watchtime for all videos and not only for entertainment content. The answers were recalculated into minutes when necessary (M = 302.16, SD = 407.38).

Perceived user agency. To measure the perceived user agency of participants over their YouTube consumption, the Sense of Agency Scale (SoAS) as proposed by Tapal et al. (2017) was implemented. The original scale contains thirteen items of which nine were deemed appropriate to use in this study. These statements were adjusted in wording to fit the research context (e.g., "The decision whether and when to watch videos is within my own hands", "I am just an instrument in the hands of the recommendation system"). The items are scored using a Likert scale ranging from 1 *(strongly disagree)* to 5 (*strongly agree),* instead of a seven-point scale like the study by Tapal et al. (2017) presented to create internal consistency. The overall level of perceived user agency was established by calculating the average score across the items (*M* = 3.38, *SD* = 0.55; Cronbach's α = .70).

Recommendation satisfaction. In order to measure how satisfied users are with their YouTube recommendations, one of the four items from the scale used by Liu et al. (2010) was adapted. They used a seven-point semantic scale (ranging from -3 to 3), asking: "My overall opinion on my video recommendations on YouTube is:" with the adapted item being "(-3) Very dissatisfied/ Very satisfied (3)". The semantic scale was changed into a five-point semantic scale for internal consistency with the answer options "(-2) Very dissatisfied/ Very satisfied (2)". The overall level of recommendation satisfaction was measured by calculating the mean score across the items (*M* = 3.75, *SD* = 0.75).

Content/social/process/technology gratifications. For the four gratifications, an existing scale with several items per gratification type was implemented by Liu et al. (2010). This scale has been used in multiple studies that focus on YouTube already (e.g., Bakar et al., 2014; Balakrishnan & Griffiths, 2017) which improves the construct validity. The statement "Compared to your expectations before using YouTube, how do you experience YouTube to perform the following functions:" was posed, followed by fifteen items, for which the answer options were “Much lower than your expectation” (-3), “Just the same as your expectation” (0) and “Much higher than your expectation” (3), on a seven-point scale. For internal consistency, this was changed into a five-point scale using the same answer options being “Much lower than your expectation” (-2), “Just the same as your expectation” (0) and “Much higher than your expectation” (2). An example of the five items from the content gratifications scale is "To provide information". An example of the two items from the social gratifications scale is "To connect with persons who share some of my values ". " I have nothing better to do " is one of the four items of the process gratifications scale. Finally, "I can use it anytime, anywhere" is an example of the four items from the technology gratifications scale. The gratification score was determined by calculating the average score per gratification (*Content: M* = 3.46, *SD* = 0.66; Cronbach's α = .57), (*Social: M* = 2.45, *SD* = 0.84; Cronbach's α = .68), (*Process: M* = 3.64, *SD* = 0.63; Cronbach's α = .76), (*Technology: M* = 3.52, *SD* = 0.60; Cronbach's α = .70).

3.5 Reliability of the measurements

In order to safeguard the internal consistency of the items that created the scales for the used variables factor analyses were conducted. When conducting a factor analysis, the sample size needs to be larger than 150 respondents, the scale needs at least three variables, the variables have to be measured on a continuous level and the variables need to be normally distributed (Pallant, 2016). In addition to this, reliability analyses were conducted with scales with two or more items to check whether the scale is actually measuring the desired construct. Since all items were adjusted in wording from the original scale these analyses were deemed necessary to ensure reliability and validity of the scales. Thus, factor and reliability analyses were conducted for the variables that met these conditions and the results are reported in the following paragraphs.

3.5.1 Perceptions of the recommendation algorithm

After reverse coding two items, all five items were entered into a confirmatory factor analysis using a Principal Component analysis expecting one component to be extracted. This one component was expected since the researcher created the question based on the categories by Bucher (2017) to measure a single construct based on either a more positive or negative view of the recommendation algorithm. The analysis resulted in a one-dimensional scale since one component has an Eigenvalue higher than 1.00 (Eigenvalue of 2.55, explaining 50.9% of the total variance), $KMO = .72$, $\chi 2$ ($N = 161$, 10) = 2018.09, $p < .001$. All items relate positively to the first component, as is visible in table 3.1 so a new variable was created. The scale has acceptable reliability, Cronbach's $\alpha = .70$. So, the scale seems to accurately measure user perceptions of the recommendation algorithm. Looking at the variable itself, a high score indicates a more positive perception of the recommendation system, whereas a low score indicates a more negative perception.

Table 3.1. Component loadings perceptions of the recommendation algorithm

	Component 1
The user profile that I suspect that YouTube uses to suggest videos to me is accurate.	.83
I feel like the recommendation algorithm keeps getting worse.	.81
I am very positive about the recommendation system.	.80
I think it's convenient that the algorithm understands my interests and suggests the right videos to me.	.68
I get frustrated or annoyed when the algorithm shows me videos that I am not interested in.	.30
Reliability	α = .70
Variance Explained	50.9%

3.5.2 Perceived recommendation content diversity

First, two items were reverse coded, whereafter all four items were entered into a confirmatory factor analysis using a Principal Component analysis expecting one component to be extracted, *KMO* = .66, $\chi 2$ (N = 161, 6) = 209.19, p < .001. A confirmatory factor analysis was used since the existing scale by Willemsen et al. (2016) that was implemented was based on one factor. The analysis demonstrated one component that has an Eigenvalue higher than one (Eigenvalue of 2.40, explaining 60.1% of the variance). All four items positively related to the first component which is visible in table 3.2, after which the final variable was computed. The scale has acceptable reliability, Cronbach's α = .77. Hence, the variable seems to accurately measure the construct. The higher the score, the more diverse the recommendation content is perceived as.

Table 3.2. Component loadings perceived recommendation content diversity

	Component 1
My list of recommended videos is varied and diverse.	.86
Most recommended videos are about the same topics.	.81
All of my recommended videos are similar to each other.	.73
My suggested videos differ a lot from each other in different aspects.	.68
Reliability	α = .77
Variance Explained	60.1%

3.5.3 Perceived user agency

After reverse coding five items, all nine items were entered into a confirmatory factor analysis using a Principal Component analysis expecting a single component to be extracted, *KMO* = .73, $\chi 2$ (N = 161, 36) = 220.49, p < .001. One component was expected since the scale that was based on the items by Tapal et al. (2017) was partly reverse coded, so the items measured whether an individual either experienced a high or low amount of agency. The analysis showed one component that has an Eigenvalue higher than one (Eigenvalue of 2.70, explaining 30.0% of the variance). All nine items positively related to the first component which is visible in table 3.3, hereafter the final variable was computed. The scale has acceptable reliability, Cronbach's α = .70. Thus, the variable seems to accurately measure the construct. The higher the score, the more agency a person seems to experience over their YouTube consumption.

Table 3.3. Component loadings perceived user agency

	Component 1
The videos I watch just happen without my intention.	.75
I feel like the videos that I choose to watch are controlled by the recommendation system.	.62
I am in full control of what I watch.	.61
The videos I eventually watch generally surprise me.	.60
No video I watch is actually voluntary.	.51
I am completely responsible for every video I end up watching.	.51
The decision whether and when to watch videos is within my own hands.	.42
Which videos I watch are planned by me from the very beginning to the very end.	.41
I am completely responsible for every video I end up watching.	.38
Reliability	α = .70
Variance Explained	30.0%

3.5.4 Content, social, process and technology gratifications

For the content, social, process and technology gratifications the scale by Liu et al. (2010) was used so four factors were expected and a confirmatory factor analysis was executed. Through a Principal Component Analysis four components were extracted, *KMO* = .83, $\chi 2$ (N = 161, 66) = 537.89, $p < .001$. The Eigenvalues and explained variances of the four components are visible in table 3.4 as well as the factor loadings. The final variables were then separately computed per gratification type.

Content gratifications. The first factor included two items that are related to information on the platform. However, poor reliability was encountered, so results stemming from this variable should be interpreted with caution (Cronbach's α = .57).

Social gratifications. The second factor contains two factors and measures the social implications of the platform. This scale shows questionable reliability, so again caution with interpretation is required (Cronbach's α = .68).

Process gratifications. The third factor included four items and related to the process of utilizing the platform. Acceptable reliability was encountered here (Cronbach's $\alpha = .76$).

Technology gratifications. The fourth and final factor contains four items and accounts for the technological affordances of the platform. The scale was found to be acceptable (Cronbach's $\alpha = .70$).

Table 3.4. Component loadings factor analysis content, social, process and technology gratifications

Item	**Content gratifications**	**Social gratifications**	**Process gratifications**	**Technology gratifications**
To provide information.	-.92			
To present info on my interests.	-.65			
To connect with persons who share some of my values.		.83		
To meet new people.		.88		
It's enjoyment.			.41	
It's entertainment.			.64	
It helps pass time.			.63	
I have nothing better to do.			.88	
It's convenient to use.				.66
I can get what I want for less effort.				.46
I can use it anytime, anywhere.				.58
It is easier to use.				.95
Reliability	α = .57	α = .68	α = .76	α = .70
Variance Explained	34.7%	12.9%	9.7%	7.5%

3.5.5 Control variables

It is useful to have insights into the demographic distribution of the represented variable, therefore the demographic variables *gender*, *age*, *educational level* and *nationality* were included in the analysis. These variables act as control variables in the regression analyses to check for confounding variables. Confounding variables describe the occurrence of a distortion of the effects between the dependent variable and the predictor (Pallant, 2016).

During the data preparation stage, the demographic variables *gender*, *educational level* and *nationality* were recoded into binary dummy variables. First of all, the *gender* dummy variable consists of two groups which are "Female" and "Not female". "Female" represents the largest group for the *gender* variable and encompasses all females. "Not female" encompasses all male respondents and the participants who selected "Non-binary" or "Prefer not to say". Secondly, the *educational level* variable is divided into the groups "Higher education" and "No higher education". The "Higher education" group includes participants who selected "University of applied sciences or similar", "University bachelor's degree" and "Graduate or professional degree" which is in accordance with the definition of higher education of the Dutch government (Ministerie van Onderwijs, Cultuur en Wetenschap, 2020). The "No higher education" group includes respondents who selected "Post-secondary vocational education or similar", "Middle school", "Primary school" or "Prefer not to say". For the variable *nationality* the two groups "Dutch" and "Non-Dutch" were established since the majority of the participants were Dutch (52.8%). All respondents who selected being a national from the Netherlands were included in the "Dutch" group and the other nationalities in the "Non-Dutch" group. The demographic variable *age* was not recoded since it is measured on a continuous scale.

3.6 Validity

According to Babbie (2011), construct, content, criterion, internal and external validity need to be present in the design of a study to improve the generalizability of the sample. To ensure the construct validity of the survey the pre-test was executed with multiple participants. As for content validity, the factor analyses ensured the correct inclusion of all items for the variable. Additionally, the scales that were adapted by Liu et al. (2010), Tapal et al. (2017) and Willemsen et al. (2016) as well as the scale based on the categories by Bucher (2017) enhance the criterion validity since they are established from existing research.

Regarding the internal validity of the study at hand, this type of validity is reinforced by the given explanation to the survey participants of what entertainment content on

YouTube entails as well as the referral to the YouTube app statistics for the weekly watchtime question. Threats to internal validity might be the occurrence of shared YouTube accounts by users or respondents who have two accounts which may then lead to an inaccurate estimate of the weekly average watch time. In addition to this, external validity relates to the generalizability of the results of the survey (Babbie, 2011) and is reassured by the online distribution of the survey on several international platforms. Nevertheless, a threat to external validity may be that the representation of different nationalities is skewed, as 52.8% of the sample has the Dutch nationality as is visible with the other represented nationalities in appendix table B2.

Chapter 4. Results

The upcoming chapter discusses the results from this study for which the aim is to explore the extent to which perceptions of the YouTube recommendation system shape user entertainment content consumption. First of all, an assumption check of the Skewness and Kurtosis prepares the variables for the conducted analyses. Then, the impact of perceptions of the recommendation algorithm on recommendation content diversity, watchtime, recommendation satisfaction and perceived user agency is inquired about. Moreover, the influence of content, social, process and technology gratifications on recommendation satisfaction has been explored. To study these variables, three multiple regression analyses and a hierarchical regression analysis were executed. Hence, this chapter presents the outcomes of the tested hypotheses.

4.1 Assumption check

Before executing the analyses it was checked whether the variables were normally distributed, which is an assumption that should be met before conducting a regression analysis (Pallant, 2016). This was especially relevant for the watchtime variable since the reported watchtimes showed great variation (M = 302.16, SD = 407.38). Thus, the Skewness and Kurtosis were checked for this variable. The Skewness being the asymmetry of the distribution and the Kurtosis measuring the "peakedness" of the distribution, which should be between -1 and 1 for a normal distribution (Pallant, 2016). Initially, the watchtime variable showed a Skewness value of 3.16 and a Kurtosis value of 15.56, which is unacceptable since the acceptable Skewness values range from - 3 to + 3 and the acceptable Kurtosis value ranges from - 10 to + 10 (Griffin & Steinbrecher, 2013). Therefore, the watchtime variable was recoded by applying the following formula *Ln (Watchtime + 1)* to the initial variable. Hereafter, the Skewness value was -.46 and the Kurtosis was -.55, which is acceptable. The other variables showed acceptable Skewness and Kurtosis values as is displayed in table 4.1.

Table 4.1 Assumption check Skewness and Kurtosis

Variable	Skewness	Kurtosis
Perceptions of the recommendation algorithm	.02	.43
Perceived recommendation content diversity	.35	-.74
Watchtime (log transformed)	-.46	-.55
Perceived user agency	.01	-.24
Recommendation satisfaction	-.70	.93
Content gratifications	.02	.30
Social gratifications	-.23	-.72
Process gratifications	-.24	.09
Technology gratifications	.29	-.01

4.2 Results of analyses

For testing H1, a multiple regression analysis was conducted since the relationship between a continuous independent variable and a continuous dependent variable had to be inquired about. Also, the four control variables were entered as independent variables into the analysis. The *perceived recommendation content diversity* score was used as a criterion and respectively the variables *perceptions of the recommendation algorithm, gender, nationality, educational level* and *age* were entered as predictors.

The multiple regression model was not found to be significant, F (5, 152) = .80, p = .548. Table 4.2 demonstrates the standardized beta weights and explained variance for the predictors. Thus, no associations between the independent and dependent variables were encountered and H1 was rejected.

Table 4.2. Standardized beta weights and R^2 of the multiple regression analysis with the perceived recommendation content diversity score as a criterion

	Model 1
Predictor	
Perceptions of the recommendation algorithm	.08
Gender	.01
Nationality	-.05
Educational level	.12
Age	-.01
	R^2 = .03
	p = .548

Note: *$p < .05$, **$p < .01$, ***$p < .001$.

To test H2, another multiple regression analysis was executed since a continuous dependent and a continuous independent variable were inquired about. The *watchtime (log transformed)* score was used as a criterion and respectively the variables *perceptions of the recommendation algorithm, gender, nationality, educational level* and *age* were entered as predictors.

The model reached significance, F (5, 151) = 4.93, $p < .001$. The predicting control variable *age* is statistically significant, (B = -.06, $p < .001$). Thus, watchtime reduces by .06 with every increasing year. The correlation between the variables is negatively medium (see table 4.3), 95% CI [-.09, -.03]. The variable *perceptions of the recommendation algorithm* and the other control variables were not significant predictors. Also, no evidence of multicollinearity is found as the VIF values vary between 1.01 and 1.04 and the tolerance varies between .95 and .99. An overview of the multiple linear regression analysis is presented in table 4.3. However, H2 is rejected since there is no correlation between recommendation algorithm perceptions and watchtime.

Table 4.3. Standardized beta weights and R^2 of the multiple regression analysis with the watchtime (log transformed) score as a criterion

	Model 1
Predictor	
Perceptions of the recommendation algorithm	.02
Gender	.08
Nationality	.14
Educational level	.10
Age	**-.33***
	$\boldsymbol{R^2 = .14}$
	$\boldsymbol{p < .001}$

Note: *$p < .05$, **$p < .01$, ***$p < .001$.

To test H3 and H5, a hierarchical regression analysis was executed since the research model shows that both *perceptions of the recommendation algorithm* and *content, social, process* and *technology gratifications* might affect *recommendation satisfaction.* All variables are measured on a continuous level. For the hierarchical regression analysis, in the first model *content, social, process and technology gratifications* were entered as predictors. In the second model *perceptions of the recommendation algorithm* was added and finally in the third model the control variables were added as predictors. In all models *recommendation satisfaction* was the criterion.

The first model that tests H5 reached significance, F (4, 154) = 6.03, $p < .001$. The predictor variable *content gratifications* is statistically significant and has a weak positive correlation with *recommendation satisfaction* (B = .22, p = .028), 95% CI [.02, .42], see table 4.4 for the Beta values. This means that the better content gratifications needs are met, the higher the recommendation satisfaction score is. Furthermore, the predictor variable *technology gratifications* is statistically significant and has a weak positive correlation with *recommendation satisfaction* (B = .30, p = .018), 95% CI [.05, .54]. This means that the better technology gratifications needs are met the higher the recommendation satisfaction score is.

Process and social gratifications have no significant correlation with recommendation satisfaction. A summary of the hierarchical analysis is portrayed in table 4.4. Additionally, no evidence of multicollinearity is found as the VIF values vary between 1.01 and 1.64 and the tolerance varies between .61 and .99. Hence, H5 is partially accepted since content and technology gratifications are significant predictors and do affect recommendation satisfaction.

The second model which answers H3, adds the perception variable to the analysis. This model reached significance and adding the new variable improved the model, $\Delta R^2 = .25$, $F(1, 152) = 61.48$, $p < .001$. The predictor variable *perceptions of the recommendation algorithm* is statistically significant and has a positive strong correlation with *recommendation satisfaction* ($B = .64$, $p < .001$), 95% CI [.48, .80], see table 4.4 for the Beta value. This means that more positive perceptions regarding the recommendation algorithm lead to higher recommendation satisfaction. The four gratification variables were not statistically significant anymore in the second model as is displayed in table 4.4. Moreover, no evidence of multicollinearity is found as the VIF values vary between 1.05 and 1.69 and the tolerance varies between .66 and .95. However, the correlation between the *perceptions of the recommendation algorithm* variable and *recommendation satisfaction* variable is so strong that it practically hides the effect of the other variables entered into the model. For this reason, the gratifications were added into the first model and perception were only added to the second model, so the gratifications variables were not affected by the perceptions variable. Thus, H3 is accepted. Lastly, the third model with added control variables was not significant, $\Delta R^2 = .01$, $F(4, 148) = .77$, $p = .549$.

Table 4.4. Hierarchical regression model with the recommendation satisfaction score as a criterion

	Model 1 beta weights	**Model 2 beta weights**	**Model 3 beta weights**
Predictor			
Content gratifications	**.19***	.10	.08
Social gratifications	.07	-.03	-.02
Process gratifications	.00	.01	-.01
Technology gratifications	**.23***	.13	.13
Perceptions of the recommendation algorithm		**.54*****	**.53*****
Gender			.06
Nationality			.08
Education			-.02
Age			-.05
R^2	.14	.39	.40
F	**6.03*****	**67.51*****	68.28
ΔR^2		.25	.01
ΔF		**61.48*****	.77

Note: $*p < .05$, $**p < .01$, $***p < .001$.

For testing H4, also a multiple regression analysis was conducted since the relationship between a continuous independent variable and a continuous dependent variable needed to be studied. Also, the four control variables were entered as independent variables into the analysis. The *perceived user agency* score was used as a criterion and respectively the variables *perceptions of the recommendation algorithm, gender, nationality, educational level* and *age* were entered as predictors.

The multiple regression model was not found to be significant F (5, 152) = 1.83, p = .111. Table 4.5 demonstrates the standardized beta weights and explained variance for the predictors. Thus, no associations between the independent and dependent variables were encountered and H4 was rejected.

Table 4.5. Standardized beta weights and R^2 of the multiple regression analysis with the perceived user agency score as a criterion

	Model 1
Predictor	
Perceptions of the recommendation algorithm	.10
Gender	-.12
Nationality	.12
Educational level	.06
Age	.13
	R^2 = .06
	p = .111

Note: *p < .05, **p < .01, ***p < .001.

4.3 Summary of accepted and rejected hypotheses

To verify the posed hypotheses that were based on the current body of academic literature three multiple regression analyses and a hierarchical regression analysis were conducted. Table 4.6 presents an overview of the accepted and rejected hypotheses.

Table 4.6. Accepted and rejected hypotheses

	Outcome
Hypothesis	
H1: Positive perceptions regarding the recommendation algorithm lead to less diverse entertainment content recommendations.	**Rejected**
H2: Positive perceptions regarding the recommendation algorithm lead to more time spent consuming entertainment content.	**Rejected**
H3: Positive perceptions regarding the recommendation algorithm lead to higher user satisfaction concerning their entertainment content recommendations.	**Accepted**
H4: Negative perceptions regarding the recommendation algorithm lead to a lower perceived amount of agency over the user's entertainment content consumption.	**Rejected**
H5a: The user's content gratifications score positively affects the user satisfaction score regarding their entertainment content recommendations.	**Accepted**
H5b: The user's social gratifications score positively affects the user satisfaction score regarding their entertainment content recommendations.	**Rejected**
H5c: The user's process gratifications score positively affects the user satisfaction score regarding their entertainment content recommendations.	**Rejected**
H5d: The user's technology gratifications score positively affects the user satisfaction score regarding their entertainment content recommendations.	**Accepted**

Chapter 5. Discussion

This study aims to inquire about the extent to which perceptions of YouTube's recommendation algorithm shape user entertainment content consumption. The following chapter discusses the most relevant findings of this research. First, the theoretical implications will be presented per hypothesis by discussing the results in the light of existing academic literature. Second, the societal implications of this research will be addressed. Thirdly, the limitations and strengths of the study are presented. Finally, suggestions for future research are provided.

5.1 Theoretical implications

5.1.1 The algorithmic shaping of perceived entertainment content diversity

The first hypothesis studied the correlation between perceptions regarding the recommendation algorithm and the diversity of the entertainment content recommendations that are offered by the system. No significant relationship was encountered for this hypothesis, which means that perceptions of the recommendation algorithm do not affect entertainment recommendation diversity on YouTube.

There might be several explanations for this outcome. First of all, even though filter bubbles were discovered among popular videos on YouTube (Roth et al., 2020), other studies show that these confinement dynamics mostly occur among videos with an extreme right-wing ideology or discussing conspiracy theories (Bryant, 2020; Goodrow, 2021). These videos are more informative rather than entertaining, hence they fall outside the scope of this study. Filter bubbles in the entertainment genre on YouTube have not been studied in a computational manner yet in current literature, which may also be because they simply do not occur or are deemed irrelevant.

Another reason why a relationship between perceptions of the recommendation system and diversity was not found may be that the sample had difficulty in determining whether their entertainment recommendations were diverse. YouTube users may be so accustomed to their own recommendations that they perceive them to be more diverse and renewing than they are, or they might actually be diverse already. In addition to this, it may be easier to notice confinement within informational or political content on YouTube, since the media more often discusses polarization, the notion in politics where left and right are becoming more oppositional (Bruns, 2019). So, this type of confinement in diversity might be easier to spot because YouTube users may be more actively aware of this.

Finally, some researchers have been critical of the effect of filter bubbles in general, since strong empirical evidence is lacking (Bruns, 2019). Bruns (2019), argues that filter bubbles stand as a convenient technological scapegoat that fuels the existing moral panic caused by political and social polarization. The author claims that this is a critical issue and the algorithmic shaping of the public only plays a minor role since polarization has societal causes that cannot be solved by adjustments in technology alone (Bruns, 2019). Hence, this research encountered that positive or negative perceptions of the algorithm do not significantly influence entertainment recommendation diversity. This outcome supports the conclusion from Bruns (2019) by finding a lack of empirical evidence for the occurrence of filter bubbles in the realm of entertaining content.

5.1.2 The algorithmic shaping of watchtime

The second hypothesis addressed the effect of perceptions of the recommendation system on the amount of watchtime users had. There was no significant relationship found between these perceptions and watchtime, which means that in this study the manner in which users viewed the recommendation algorithm did not lead to a higher or lower watchtime specifically.

This is an interesting outcome, since one might suspect that a user who for instance is more negative about the recommendation system and thinks it is intrusive rather than convenient might watch a lot less YouTube videos than someone who has a more neutral or positive opinion. This phenomenon might be explained by the privacy paradox theory (Kokolakis, 2017), which argues that individuals may claim to care about their privacy and personal information when you ask about it, yet their online behavior may show otherwise. Often people, especially young adults, are found to display a critical attitude regarding their privacy and they are understanding of the possible risks that come with disclosing personal information online (Hargittai & Marwick, 2016). However, simultaneously they provide their personal data to organizations without complete awareness of the consequences of doing so. Thus, YouTube users who may have negative perceptions of the recommendation algorithm might still be inclined to watch many videos and hence share their data and interests with Google as well as third parties who provide advertisements that are visible on unpaid YouTube accounts. So, the awareness of possible privacy infringements does not outweigh the high entertainment factor of the platform and this finding supports the notion of the privacy paradox theory.

Another explanation for this finding might be that the algorithm is so well designed and tailored to the user's needs that users are easily convinced to continue watching, whether they like the recommendation system or not. This is in line with the intended purpose of the algorithm, which is a success rate that is based on increased watchtime (Covington et al., 2016). Interestingly, age was found to be a significant predictor of watchtime, in the sense that younger users had a higher watchtime. This corresponds with a study by Klobas et al. (2018) among students of which almost 20% reported that they experienced compulsive usage and lack of control over their consumption. Hence, it seems that young people are relatively more susceptible to watch YouTube videos for longer periods of time and this outcome substantiates the results from the study by Klobas et al. (2018).

5.1.3 The algorithmic shaping of recommendation satisfaction

The third hypothesis concerned the relationship between perceptions of the recommendation system and user satisfaction concerning their entertainment content recommendations. A significant relationship was encountered between the two variables, which means that more positive perceptions of the recommendation algorithm led to higher user satisfaction with their entertaining recommendations.

This phenomenon is likely to be the case because users who hold a more positive attitude regarding the recommendation system may trust the algorithmic process that selects their entertainment recommendations for them better. This has been conceptualized in literature as algorithm appreciation, where people prefer the judgment of AI over their own judgment (Logg et al., 2019). Hence, rather than actively searching for new videos to watch by entering queries, users who think more highly of the algorithm simply let the system do its job and provide them with entertainment recommendations which are satisfying to them. The high level of user specificity from the deep learning algorithm is probably what has led to this satisfaction with the recommendations. This means that when the user has positive experiences with AI, this leads to more trust in the products that the algorithm offers, which are the recommendations. Shin et al. (2020) found a very similar result in their research that studied perceptions of algorithmic features like transparency and fairness and how these factors influence user satisfaction and emotion with regards to AI technologies. Hence, from the research at hand can be derived that positive perceptions of the YouTube algorithm are linked to higher user satisfaction regarding entertainment recommendations and this finding supports the conclusions of previous research.

5.1.4 The algorithmic shaping of perceived user agency

The fourth hypothesis inquired about the relationship between perceptions regarding the recommendation algorithm and the perceived amount of agency a user has over their entertainment content consumption. No significant result was found for this hypothesis, meaning that a positive or negative perception of the recommendation system does not influence the degree to which a user feels that they exert agency over their content consumption.

This hypothesis was created with the U&G theory by Katz et al. (1973-1974) in mind. Other academics also let their work regarding the YouTube platform be inspired by this theory (e.g., Bakar et al., 2014; Balakrishnan & Griffiths, 2017; Khan, 2017), however no relationship between the studied variables was found for this research. This might be the case because the U&G theory focused on print media, radio and television as the primary media sources and it therefore might be relatively outdated (Katz et al., 1973-1974). The U&G theory was coined in a time before social media and AI technologies which might reduce the applicability of the principles of the theory. Perhaps, people find being an active agent in making media consumption choices and being fully aware of the reasoning behind these choices of lesser importance these days. This might actually be caused by the fact that most users have some form of algorithmic awareness when they use social media platforms and they therefore simply assume that their agency is taken away to a certain extent (Bucher, 2017).

This also relates to the concept of the Human-AI interaction approach as discussed by Sundar (2020), where users view the platform as intelligent and capable of modifying content in a revolutionary manner. So, when using a platform like YouTube, the user is aware that they are interacting with AI and that that might alter their online experience. This is something a user might get used to and therefore they might be less affected or even be indifferent of their declination of agency over their entertainment consumption habits.

Moreover, the entertainment genre may be seen as a relatively low stakes subject as opposed to health or finances. So, users might be more prone to give away more agency to the algorithm since the consequences of doing so are less invasive. Hence, the manner in which YouTube users perceive the recommendation system was not found to affect their perceived amount of agency over their entertainment content consumption.

5.1.5 The influence of algorithmic imaginary

Next, an additional critical note on the concept of algorithmic imaginary in the context of YouTube recommendations will be provided. This is because this study discovered that the perceptions that people hold about the recommendation algorithm affect certain aspects of content consumption and not others. Since only one out of the four hypotheses about perceptions of the platform was significant, this might mean that the actual effects of the algorithmic imaginary as coined by Bucher (2017) might not be as prevalent as is thought of. Hence, the algorithmic imaginary that users hold of the AI technologies behind the platform might be existing, yet it is debatable whether entertainment content consumption is affected by these perceptions. Additionally, Bucher (2017) executed her research by conducting interviews, which is a research method that is less generalizable for the population than quantitative studies with larger samples. So, since this study is more fit for generalizable results, these results imply that in society YouTube users are not as influenced by their perceptions of the recommendation algorithm as researchers might suppose.

5.1.6 The relationship between user gratifications and recommendation satisfaction

The fifth and final hypothesis concerned the relationship between content, social process and technology gratifications and user satisfaction regarding their entertainment content recommendations. Here, content and technology gratifications were found to significantly affect recommendation satisfaction. Social and process gratifications on the other hand were not significant predictors. First the significant gratifications will be discussed and thereafter the insignificant gratifications are addressed.

The needs that are met when the content gratification score is high are those that concern information provision of the platform. When a user is satisfied with the information provision they receive when watching YouTube videos, the user becomes more satisfied with their entertainment content recommendations according to the result. This might be due to existing satisfaction with the information from the videos that is then reinforced by the tailored recommendations that may give additional information on certain subjects. However, the content gratification scale has poor reliability, so this result should be interpreted with caution. As for technology gratifications, when the needs for convenience of using the platform are met, this leads to higher satisfaction with the user's recommendations. This might be explained by the idea that users who are satisfied with the online environment that is created on the YouTube platform may also be satisfied with their recommendations that are a part of this environment and the UX. The UI has been possibly designed with technology gratifications in mind and recommendations have been conveniently placed below videos that are playing and at the home screen, which makes them easy to find and if properly suggested by the algorithm, useful to the user.

Regarding social gratifications, users may have the need to socially connect with other users. This variable did not significantly affect recommendation satisfaction, which is possibly related to the fact that YouTube is a social media platform that is mostly focused on direct entertainment and information provision and not as much on facilitating social interactions (Arthurs et al., 2018). Users do have the opportunity to comment on videos and on each other's comments, however who the person behind the profile is, is often unclear since there are no rich and publicly accessible user profiles like Instagram has for example. So, it is comprehensible that social gratifications do not affect how satisfied users are with their entertainment recommendations. Hence, social gratifications are not as applicable to YouTube as they are to other social networking sites. Yet, the social gratifications scale has questionable reliability, so the interpretation of this result needs to be done with critical awareness.

At last, process gratifications entail the needs that the platform meets when it provides entertainment, enjoyment and a form of passing time. It seems that YouTube does provide these processes to their users, however it does not affect how satisfied users are with their entertainment recommendations. This might be the case because the practical process of using the medium steers clear from the content provision that recommendations provide, so process gratifications are not applicable to recommendation satisfaction.

5.2 Societal implications

Concerning the societal implications of this study that looks at algorithms from a social embeddedness perspective rather than a computational one, multiple points can be made.

It is worth noting again that the YouTube platform is used by millions of people per day and that they consequently also interact with the recommendation system (Statista, 2021). So, since the platform is so ubiquitously used, especially in the Western world, it seems crucial to understand how the recommender system might potentially influence user consumption. It is without doubt that the YouTube algorithm became utterly successful at making the right video content suggestions, since 70% of the watched videos on the platform stem from recommendations (Nicas, 2020). However, the results of this study show that the effects of perceptions of the recommender system on entertainment content consumption are only limited, so on the individual level users might not portray altered consumer behavior that is caused by their perceptions of the algorithm. Hence, the hype around AI technologies influencing user behavior in academia, popular culture, the business world of digital media and legislation might therefore be ungrounded and this should be investigated further. It should however be made explicitly that this study only took entertainment content consumption into consideration without taking political and informational content into account.

Moreover, the results and the scope of this study discuss the future of entertainment content distribution within a large part of society. Hence, it might be interesting to acknowledge that the recommendation system as a key affordance of the platform in some way shapes content consumption. Media platforms that guide users by implementing AI technologies are thought to be the future, so this study only highlights and further investigates this argument. The findings of this study may also apply to user experiences with other platforms such as Netflix or Spotify. Like YouTube, these platforms also use recommender systems that are very much personalized to the user's unique taste and they suggest their content in similar ways as well. So, the results of this research may also be applicable to other digital media platforms and create higher awareness of the possible effects of algorithmic imaginary. This awareness might aid these companies with further developments of their recommendation algorithm, UI or UX using the presented empirical evidence.

Also, the finding that positive perceptions of the algorithm as well as technology gratifications needs that are met increase recommendation satisfaction might be especially useful to the digital media business world. This result substantiates how valuable proprietary recommender systems are to a company and that algorithms that operate successfully actually create higher content satisfaction. This as a consequence is likely to lead to increased content consumption, which as a result leads to higher profit, the main focus and preferred outcome for most businesses.

Another societal implication of this study is that it may alert governments and other social actors or institutions of the finding that especially youngsters have the tendency to have increased watchtimes. Other studies already encountered compulsive usage (Klobas et al., 2018) and addiction to YouTube (Balakrishnan & Griffiths, 2017) among their samples, so it is important to be aware of these tendencies of over usage for young people. This way, possible mental health issues or other social problems might be evaded.

5.3 Limitations and strengths

As every other study, this research comes with limitations, however also with certain strengths.

To begin with the limitations, what is important to consider is that many of the measurements that were used in the survey were based on the objective indications of the participants' YouTube consumption habits and opinions on their recommendations and the platform. This type of measurement is however sometimes lacking in accuracy since users might be unable to objectively determine what their habits are which may have led to a decline in validity. Also, the survey was held at one moment in time, whereas perceptions of the algorithm for example may fluctuate due to news coverage or other events.

Furthermore, two variables that were used in the analysis showed poor reliability, which means that the internal consistency of these scales was not completely up to standard. Another limitation might be that even though the survey clarified that participants had to only take their entertainment content consumption and recommendations into account, for some this might have been too difficult to fully distinguish whilst answering the questions. Also, the way in which recommendations are implemented into the UI is different on for example a smart-tv than it is on a mobile phone, which may have led to differences in the perceptions of recommendations.

However, as mentioned before this research also has its strengths. A strength being the adequate size of the sample data on which the analyses are based. Also, for the watchtime variable participants were encouraged to verify their weekly watchtime in the YouTube app, which might have led to a more accurate number rather than when the researcher would have only relied on self-estimates. Finally, the new scale that was created for this study as well as the adjusted scales from existing studies are a just contribution to the existing academic literature on the discussed topic.

5.4 Suggestions for future research

Following, a few suggestions for future research that may be of interest will be addressed. First of all, it might be helpful to execute a computational and machine learning driven study on the same subject. For example, the path that entertainment content recommendations follow within the recommendation algorithm might be inquired about, which is similar to the work by Roth et al. (2020).

Another element of YouTube that might be necessary to explore is how users utilize organic consumption from search queries as opposed to algorithmic consumption which is caused by the recommender system. This might actually affect content diversity as was discovered in a similar study by Anderson et al. (2020) on the effects of the Spotify algorithm.

In addition to this, it might be interesting to study the algorithmic experience more in the context of the UI and how the placement and salience of recommendations affect consumption habits. The notion of push notifications with recommendations for new videos to watch and their effects is also still untrodden ground in the academic world.

At last, the study at hand might be improved if it is repeated overtime using the same respondents, thus transforming it into a longitudinal study. This might be interesting since consumption habits may change over time as well as opinions and perceptions. The platform YouTube itself is clearly also constantly updated by its engineers which may lead to different user experiences and attitudes in the long run. Moreover, this quantitative study could also be transformed into a qualitative study that seeks for a more in-depth understanding of user experiences with entertaining YouTube recommendations.

Chapter 6. Conclusion

The research at hand will be finalized in this chapter by providing concluding remarks on the object of study which is the YouTube recommender system and how user perceptions affect entertainment consumption. Studies that discuss the YouTube recommender system are existent, however they are not numerous. In addition to this, studies that regarded user perceptions of algorithms, termed as algorithmic awareness by Bucher (2017), were often executed in a qualitative matter (e.g., Alvarado et al., 2020; Bishop, 2019; Bucher, 2017). Furthermore, the recommendation algorithm is often discussed in literature in relation to political extremist content, whereas this study specifically focuses on entertainment content consumption which is a novel stance on the subject matter. So, the field of research on which the posed research question is based is relatively novel, and the aim of this study is to explore this untrodden ground and contribute to the existing body of works.

The following research question created the core of this study:

To what extent do perceptions of YouTube's recommendation algorithm shape user entertainment content consumption?

This question was answered using five hypotheses and a quantitative research method, being a survey. It was discovered that positive perceptions regarding the recommendation algorithm do not affect perceived diversity in entertainment content recommendations, nor influence watchtime and neither affect the perceived amount of agency over the user's entertainment content consumption. A significant effect was encountered between perceptions regarding the recommendation algorithm and entertainment recommendation satisfaction, where more positive perceptions lead to higher recommendation satisfaction. Finally, the fifth hypothesis inquired about user gratifications and whether these affect recommendation satisfaction. Here, social and process gratifications had no significant effect, yet both content and technology gratifications were positively related to recommendation satisfaction.

From these findings, three outcomes stood out in particular. The first key takeaway from this study is that perceptions of the recommender system were not found to influence watchtime. This is an interesting result since it is apparently unimportant what type of attitude a user holds against the algorithm, since this does not make them watch more or less videos. Thus, users who have a negative attitude, do not necessarily watch fewer videos. This phenomenon might be explained by the privacy paradox theory, which argues that media users might reason that they care a lot about their privacy and data protection, whilst they may simultaneously share their personal data with companies to make use of their services (Kokolakis, 2017). So, for YouTube entertainment content, the awareness of possible privacy infringements does not outweigh the attractiveness of the high entertainment factor of the platform.

The second key takeaway concerns the finding that positive perceptions regarding the algorithm do lead to more satisfaction with entertainment recommendations. This finding substantiates the algorithm appreciation theory. This theory discusses the phenomenon where people prefer the judgment of AI over their own judgment (Logg et al., 2019). Hence, when users have a positive experience with the YouTube algorithm, they also have more trust in its products, which are the recommendations. This might be an especially interesting finding for digital media companies that utilize proprietary recommender systems, since it justifies the importance of creating an optimized algorithm as well as a refined UI and UX to increase user satisfaction and consequentially consumption which eventually boosts advertisement revenue.

The third key takeaway concerns the finding that when the needs for technology gratifications are met by the YouTube platform, users are also more satisfied with their entertainment recommendations. Meaning that when users experience convenience whilst browsing through YouTube, they have higher user satisfaction which again may lead to increased usage and increased economic benefits for the company. This outcome might again incentivize digital media companies to improve their algorithm and UI as much as possible to create a better UX, since this only has positive effects on recommendation satisfaction.

On a final note, Bucher (2017) concludes her study by stating that it is crucial to understand the emotions that come with using AI in order to understand its social power. Nonetheless, this study shows that the effect of the algorithmic imaginary might be debatable since only few significant effects were encountered. Therefore, the hype in social sciences, popular culture, the business world and political affairs surrounding the influence of AI technologies on users might perhaps be overdue. However, this is only a single study and more investigation is necessary to reinforce and substantiate this idea further. Thus, the YouTube recommender system undeniably shapes entertainment content consumption in certain ways, just perhaps not the ways we suspected it to.

Reference list

Abbas, M., Riaz, M. U., Rauf, A., Khan, M. T., & Khalid, S. (2017, December). Context-aware Youtube recommender system. *In 2017 international conference on information and communication technologies (ICICT)* (pp. 161-164). IEEE. https://doi.org/10.3390/math7080740

Alvarado, O., Heuer, H., Vanden Abeele, V., Breiter, A., & Verbert, K. (2020). Middle-Aged Video Consumers' Beliefs About Algorithmic Recommendations on YouTube. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW2), 1-24. https://doi.org/10.1145/3415192

Alvarado, O., & Waern, A. (2018, April). Towards algorithmic experience: Initial efforts for social media contexts. In *Proceedings of the 2018 chi conference on human factors in computing systems, 286,* (pp. 1-12). https://doi.org/10.1145/3173574.3173860

Anderson, A., Maystre, L., Anderson, I., Mehrotra, R., & Lalmas, M. (2020, April). Algorithmic effects on the diversity of consumption on spotify. In *Proceedings of The Web Conference 2020* (pp. 2155-2165). https://doi.org/10.1145/3366423.3380281

Arthurs, J., Drakopoulou, S., & Gandini, A. (2018). Researching youtube. *Convergence, 24*(1), 3-15. doi: 10.1177/1354856517737222

Babbie, E. (2011). *Introduction to social research* (5th ed.). Thomson Wadsworth.

Bakar, M. S. A., Bolong, J., Bidin, R., & Mailin, B. (2014). Factors of gratification contributing in continuance intention to watch movies on YouTube. *Procedia-Social and Behavioral Sciences*, *155*, 9-13. doi: 10.1016/j.sbspro.2014.10.248

Bandura, A. (2005). The evolution of social cognitive theory. In K. G. Smuth & M. A. Hitt (Eds.) *Great minds in management,* (pp. 9-35). Oxford: Oxford University Press.

Banker, S., & Khetani, S. (2019). Algorithm overdependence: How the use of algorithmic recommendation systems can increase risks to consumer well-being. J*ournal of Public Policy & Marketing*, *38*(4), 500-515. https://doi.org/10.1177%2F0743915619858057

Balakrishnan, J. & Griffiths, M. D. (2017). Social media addiction: What is the role of content in YouTube? *Journal of behavioral addictions*, *6*(3), 364-377. https://doi.org/10.1556/2006.6.2017.058

Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S., Ravichandran, D. & Aly, M. (2008). Video suggestion and discovery for youtube: taking random walks through the view graph. *In Proceedings of the 17th international conference on World Wide Web,* 895-904. https://doi.org/10.1145/1367497.1367618

Bishop, S. (2019). Managing visibility on YouTube through algorithmic gossip. *New media & society, 21*(11-12), 2589-2606. https://doi.org/10.1177/1461444819854731

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77-101. http://dx.doi.org/10.1191/1478088706qp063oa

Bruns, A. (2019). It's not the technology, stupid: How the 'Echo Chamber'and 'Filter Bubble'metaphors have failed us. *International Association for Media and Communication Research.* https://eprints.qut.edu.au/131675/

Bryant, L. V. (2020). The YouTube algorithm and the Alt-Right filter bubble. *Open Information Science, 4*(1), 85-90. https://doi.org/10.1515/opis-2020-0007

Bucher, T. (2017). The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, communication & society, 20*(1), 30-44. https://doi.org/10.1080/1369118X.2016.1154086

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research, 56*(5), 809-825. https://doi.org/10.1177%2F0022243719851788

Chau, C. (2010). YouTube as a participatory culture. *New directions for youth development, 2010*(128), 65-74. https://doi.org/10.1002/yd.376

Covington, P., Adams, J., & Sargin, E. (2016, September). Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems* (pp. 191-198). https://doi.org/10.1145/2959100.2959190

Cunningham, S., & Craig, D. (2017). Being 'really real'on YouTube: authenticity, community and brand culture in social media entertainment. *Media International Australia, 164*(1), 71-81. https://doi.org/10.1177%2F1329878X17709098

Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., ... & Sampath, D. (2010, September). The YouTube video recommendation system. In *Proceedings of the fourth ACM conference on Recommender systems* (pp. 293-296). https://dl.acm.org/doi/proceedings/10.1145/1864708

DeVito, M. A., Gergle, D., & Birnholtz, J. (2017). "Algorithms ruin everything": #RIPTwitter, folk theories, and resistance to algorithmic change in social media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp.3163–3174). ACM. https://doi.org/10.1145/3025453.3025659

Dietvorst, B., Simmons, J. P., & Massey, C. (2015). Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err. *Journal of Experimental Psychology: General, 144*(1), 114-126. http://dx.doi.org/10.1037/xge0000033

Ferwerda, B., Graus, M. P., Vall, A., Tkalcic, M., & Schedl, M. (2017, April). How item discovery enabled by diversity leads to increased recommendation list attractiveness. In *Proceedings of the Symposium on Applied Computing* (pp. 1693-1696). http://dx.doi.org/10.1145/3019612.3019899

Goodrow, C. (2021, September 15). *On YouTube's recommendation system*. Blog.Youtube. Retrieved April 4, 2022, from https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/

Griffin, M. M., & Steinbrecher, T. D. (2013). Large-scale datasets in special education research. In *International Review of Research in Developmental Disabilities* (Vol. 45, pp. 155-183). Academic Press.

Hargittai, E., & Marwick, A. (2016). "What can I really do?" Explaining the privacy paradox with online apathy. *International journal of communication*, *10*, 1-21. https://ijoc.org/index.php/ijoc/article/view/4655

Haroon, M., Chhabra, A., Liu, X., Mohapatra, P., Shafiq, Z., & Wojcieszak, M. (2022). YouTube, The Great Radicalizer? Auditing and Mitigating Ideological Biases in YouTube Recommendations. *Computers and Society*. 1-21. https://doi.org/10.48550/arXiv.2203.10666

Katz, E., Blumler, J. G., & Gurevitch, M. (1973-1974). Uses and gratifications research. *The public opinion quarterly*, *37*(4), 509-523. http://www.jstor.org/stable/2747854

Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in human behavior, 66,* 236-247. https://doi.org/10.1016/j.chb.2016.09.024

Klobas J.E., McGill T.J., Moghavvemi S. & Paramanathan T. (2018). Compulsive YouTube usage: A comparison of use motivation and personality effects. *Computers in Human Behavior, 87*, 129-139. https://doi.org/10.1016/j.chb.2018.05.038

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security, 64,* 122-134. https://doi.org/10.1016/j.cose.2015.07.002

Kuss, D. J., & Griffiths, M. D. (2017). Social networking sites and addiction: Ten lessons learned. *International Journal of Environmental Research and Public Health, 14*(3), 311-328. doi: 10.3390/ijerph14030311

Lee, S. Y., & Lee, S. W. (2022). Normative or Effective? The Role of News Diversity and Trust in News Recommendation Services. *International Journal of Human–Computer Interaction,* 1-14. https://doi.org/10.1080/10447318.2022.2057116

Liu, I. L., Cheung, C. M., & Lee, M. K. (2010). Understanding Twitter usage: What drive people continue to tweet. *Proceedings from Pacific Asia Conference on Information Systems (PACIS).* Taipei, Taiwan: PACIS. https://doi.org/10.1016/j.cose.2015.07.002

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes, 151*, 90-103. https://doi.org/10.1016/j.obhdp.2018.12.005

Ministerie van Onderwijs, Cultuur en Wetenschap. (2020, June 24). Higher education. Secondary Vocational Education (MBO) and Higher Education | Government.Nl. Retrieved May 20, 2022, from https://www.government.nl/topics/secondary-vocational-education-mbo-and-higher-education/higher-education

Pallant, J. (2016). *SPSS Survival Manual* (6th ed.). Routledge.

Möller, A. M., Kühne, R., Baumgartner, S. E., & Peter, J. (2019). Exploring user responses to entertainment and political videos: An automated content analysis of YouTube. *Social Science Computer Review, 37*(4), 510-528. https://doi.org/10.1177%2F0894439318779336

Morgan, K. (n.d.). Are people completing surveys on mobile devices? SurveyMonkey. Retrieved June 19, 2022, from https://www.surveymonkey.com/curiosity/are-people-completing-surveys-on-mobile-devices/

Munger, K., & Phillips, J. (2022). Right-wing YouTube: a supply and demand perspective. *The International Journal of Press/Politics, 27*(1), 186-219. https://doi.org/10.1177/1940161220964767

Nicas, J. (2020, July 10). *Can YouTube Quiet Its Conspiracy Theorists?* The New York Times. https://www.nytimes.com/interactive/2020/03/02/technology/youtube-conspiracy-theory.html

Nwabueze, C., & Okonkwo, E. (2018). Rethinking the bullet theory in the digital age. *International Journal of Media, Journalism and Mass Communications, 4*(2), 1-10. http://dx.doi.org/10.20431/2454-9479.0402001

O'Donovan, C., Warzel, C., McDonald, L., Clifton, B., & Woolf, M. (2019, January 24). *We followed YouTube's recommendation algorithm down the rabbit hole.* Retrieved

March 27, 2022, from BuzzFeed News website: https://www.buzzfeednews.com/article/carolineodonovan/down-youtubes-recommendation-rabbithole

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think.* Penguin books.

Potter, W. J. (2014). A critical analysis of cultivation theory. *Journal of communication, 64*(6), 1015-1036. https://doi.org/10.1111/jcom.12128

Roth, C., Mazières, A. & Menezes, T. (2020). Tubes and bubbles topological confinement of YouTube recommendations. *PLoS ONE, 15*(4): e0231703. https://doi.org/10.1371/journal.pone.0231703

Salkind, N. J. (2010). *Encyclopedia of research design* (Vols. 1-0). SAGE Publications, Inc. doi: 10.4135/9781412961288

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In *The adaptive web* (pp. 291-324). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-72079-9_9

Schellewald, A. (2022). Theorizing “Stories About Algorithms” as a Mechanism in the Formation and Maintenance of Algorithmic Imaginaries. *Social Media+ Society, 8*(1), doi: 20563051221077025.

Shin, D., Zaid, B., & Ibahrine, M. (2020, November). Algorithm appreciation: Algorithmic performance, developmental processes, and user interactions. In *2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)* (pp. 1-5). IEEE. https://doi.org/10.1109/CCCI49893.2020.9256470

Soland, J., Kuhfeld, M. & Rios, J. (2021). Comparing different response time threshold setting methods to detect low effort on a large-scale assessment. *Large-scale Assess Educ 9*(8), 1-21. https://doi.org/10.1186/s40536-021-00100-w

Statista. (2021). *Global social networks ranked by number of users 2021.* Retrieved January 9, 2022, from https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

Statista. (2022, March 22). *Leading websites worldwide 2021, by monthly visits.* Retrieved June 2, 2022, from https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/

Suciu, P. (2021, December 10). YouTube Removed “Dislikes” Button – It Could Impact “How To” And “Crafts” Videos. Forbes. Retrieved January 31, 2022, from https://www.forbes.com/sites/petersuciu/2021/11/24/youtube-removed-dislikes-button--it-could-impact-how-to-and-crafts-videos/?sh=69d3b9665a53

Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAII). J*ournal of Computer-Mediated Communication, 25*(1), 74-88. doi: 10.1093/jcmc/zmz026

Sundar, S. S., & Marathe, S. S. (2010). Personalization vs. customization: The importance of agency, privacy and power usage. *Human Communication Research, 36*, 298–322. doi: 10.1111/j.1468-2958.2010.01377.x.

Tapal, A., Oren, E., Dar, R., & Eitam, B. (2017). The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment. *Frontiers in psychology, 8,* 1552. https://doi.org/10.3389/fpsyg.2017.01552

Willemsen, M. C., Graus, M. P., & Knijnenburg, B. P. (2016). Understanding the role of latent feature diversification on choice difficulty and satisfaction. *User Modeling and User-Adapted Interaction*, *26*(4), 347-389. https://doi.org/10.1007/s11257-016-9178-6

Appendix A. Survey

Dear participant,

Thank you very much for your interest in this research.

I am writing my master thesis about the impact of the YouTube recommendation system. This study is intended for people aged 18 years and older and stems from the Master's programme Digitalisation, Surveillance and Societies at Erasmus University Rotterdam.

Filling out the survey will only cost you 5 minutes and it would help me a lot. The survey is anonymous and your answers will be handled confidentially. Also, participation is voluntary and you may choose to stop at any moment by closing the survey.

Thank you in advance for your time and help!

Kind regards,
Cindy

P.S.: This survey contains a completion code for SurveySwap.io

Q2 - consent Informed consent
I read the above information and I agree with participating in the survey.
I read the above information and I do not agree with participating. I would like to end the survey here.

This study is interested in the way you use YouTube for watching entertainment content videos. Entertainment is described as videos in the categories music, films, how-to and style, gaming, shows, comedy, pets and animals, people and blogs, travel, and sports.

Q4 Did you use YouTube in the past month to watch an entertaining video?
Yes
No

Q5 - age What is your age? (in years)

____

The next question asks for your weekly watchtime.
Need an indication of your watchtime? Open the YouTube app, click on your profile in the right top corner on the homepage and then click watchtime (NL: Kijktijd) to discover your consumption habits.

Q7 How much time do you spend watching entertainment videos per week?

____

Now, we're going to dive into the YouTube recommendation system/algorithm. When you open YouTube you find video recommendations on the homepage and under or next to the video that you are watching.

The following questions are about this system and the videos you watch in the entertainment genre.

Q9 Rate your level of agreement for each statement
Strongly disagree - Disagree - Neutral - Agree - Strongly agree

I am in full control of what I watch
The videos I watch just happen without my intention
The videos I watch don't logically follow from my actions
The videos I eventually watch generally surprise me
The decision whether and when to watch videos is within my own hands
Which videos I watch is planned by me from the very beginning to the very end
I feel like the videos that I choose to watch are controlled by the recommendation system
No video I watch is actually voluntary
I am completely responsible for every video I end up watching

Q10 Rate your level of agreement for each statement

Strongly disagree - Disagree - Neutral - Agree - Strongly agree

My list of recommended videos is varied and diverse
All of my recommended videos are similar to each other
Most of my recommended videos are about the same topics
The types of entertainment videos that are recommended to me have a lot of variation in their content

Q11 My overall opinion on my video recommendations on YouTube is:

1 Very dissatisfied
2
3
4
5 Very satisfied

Q12 Rate your level of agreement for each statement

Strongly disagree - Disagree - Neutral - Agree - Strongly agree

I think it's convenient that the recommendation algorithm understands my interests and recommends the right videos to me
I am aware that the recommendation algorithm tracks my video watching behavior
The user profile that I suspect YouTube uses to recommend videos to me is accurate
I get frustrated or annoyed when the recommendation algorithm shows me videos that I am not interested in
I feel like the recommendation algorithm keeps getting worse
I am very positive about the recommendation system

Q13 You're almost done, the last question about YouTube is about your experience with the platform in general. Compared to your expectations before using YouTube, how do you experience YouTube in performing the following functions:

Much lower than my expectation - Lower than my expectation - Just the same as my expectation - Higher than my expectation - Much higher than my expectation

To provide information
To share information that is useful to other people
To present info on my interests
To keep track of what I'm doing
To keep a record of what I learn
To connect with persons who share some of my values
To meet new people
It's enjoyment
It's entertainment
It helps pass time
I have nothing better to do
It's convenient to use
I can get what I want for less effort
I can use it anytime, anywhere
It is easier to use

These are the final questions!

Q15 Which gender do you identify with?
Male
Female
Non-binary
Prefer not to say

Q16 What is the highest level of education you have completed?

Primary school

Middle school

Post-secondary vocational education or similar

University of applied sciences or similar (HBO)

University Bachelor's degree

Graduate or professional degree

Prefer not to say

Q17 - nationality Which country are you a national of?

▼ Netherlands ... Zimbabwe

Q18 Which devices do you use to watch YouTube on? (Multiple answers possible)

Smartphone

Laptop/desktop

Tablet

TV screen

Other ____

Q19 Do you also create content on YouTube yourself on a regular basis? So, do you upload videos on YouTube?

Yes

No

Appendix B. Descriptive statistics

Table B1. Descriptive statistics gender and highest educational level with N = 160

Variable	Number of participants	Percentage of participants
Gender		
Female	104	65.0%
Male	51	31.9%
Non-binary	5	3.1%
Highest completed education		
Primary school	1	0.6%
Middle school	16	10.0%
Post-secondary vocational education or similar	23	14.3%
University of applied sciences or similar	31	19.4%
University Bachelor's degree	64	40.0%
Graduate or professional degree	24	15.0%
Prefer not to say	1	0.6%

Table B2. Descriptive statistics nationality with *N* = 159

Variable	**Number of participants**	**Percentage of participants**
Nationality		
Netherlands	84	52.8%
United Kingdom of Great Britain and Northern Ireland	12	7.5%
United States of America	11	6.9%
Germany	7	4.4%
Belgium	6	3.8%
Hungary	4	2.5%
Turkey	4	2.5%
Bulgaria	3	1.9%
China	3	1.9%
Philippines	3	1.9%
Brazil	2	1.2%
Canada	2	1.3%
India	2	1.3%
Malaysia	2	1.3%
Czech Republic	1	0.6%
France	1	0.6%
Hong Kong (S. A. R.)	1	0.6%
Iraq	1	0.6%
Ireland	1	0.6%
Lithuania	1	0.6%

Luxembourg	1	0.6%
Mexico	1	0.6%
Poland	1	0.6%
Portugal	1	0.6%
Singapore	1	0.6%
South Africa	1	0.6%
Ukraine	1	0.6%
Viet Nam	1	0.6%

Table B3. Descriptive statistics content creation and device usage with N = 161

Variable	**Number of participants**	**Percentage of participants**
Do you create content on YouTube?		
Yes	16	9.9%
No	145	90.1%
Which devices do you use to watch YouTube on?		
Smartphone	150	93.2%
Laptop/desktop	126	78.3%
TV screen	63	39.1%
Tablet	28	17.4%
Other: Game console(s)	1	0.6%
Other: VR	1	0.6%
Other: Smartwatch	1	0.6%