Does the Netflix recommender system produce customer utility?

An analysis of the technology acceptance of the algorithmic-prediction-based Netflix recommender system and its drivers

Daniel Lengyel

Dissertation written under the supervision of professor Boris Durisin

Dissertation submitted in partial fulfilment of requirements for the International MSc in Management with specialization in Entrepreneurship & Innovation, at the Universidade Católica Portuguesa and the ESCP Europe Paris, June 2021.

Abstract

Title: Does the Netflix recommender system produce customer utility? An analysis of the technology acceptance of the algorithmic-prediction-based Netflix recommender system and its drivers

Author: Daniel Lengyel

Key words: Technology acceptance model, recommender system, utility, algorithm, streaming, surveillance capitalism, platform, Netflix

In the last two decades technology companies engaging in surveillance capitalism (gathering data, creating predictive products, implementing behavioural modification) have reaped exorbitant profits. Using big data and machine learning algorithmic recommender systems are able to accurately predict future behaviour of users. However, other aspects than accuracy should be considered for the success of such systems. From a company perspective, Netflix has been successfully engaging in surveillance capitalism in video streaming. Claiming to be a user-centric company, personalised recommendations are the basis of Netflix's success, while it ventures into new strategic directions with original content. Using the technology acceptance model to adopt a user-perspective, this paper examines the utility of users of the Netflix recommender system. The effects of transparency and curation as features of the Netflix platform together with user's level of trust towards the system are examined for their impact on the perceived usefulness and ease of use of the recommender system, in order to determine user's behavioural intent to use the system and actual system usage. Additionally, the potential effects of user interaction as a potential future feature are explored. Using structural equation modelling on data collected from survey respondents, the paper finds that curation and trust in fact impact behavioural intent and usage while transparency fails to impact perceived factors. User interaction does not significantly improve the utility of users. The outcome suggests that Netflix should focus on curation and trust-building features as differentiating characteristics of their platform to sustain competitive advantage.

Sumário

Titulo: O sistema de recomendação Netflix produz utilidade para o cliente? Uma análise da aceitação da tecnologia do sistema Netflix recommendender baseado na previsão algorítmica e os seus drivers

Autor: Daniel Lengyel

Palavras-chave: Modelo de aceitação de tecnologia, sistema de recomendação, utilidade, algoritmo, streaming, capitalismo de vigilância, plataforma, Netflix

Nas últimas duas décadas, as empresas tecnológicas envolvidas no capitalismo de vigilância têm colhido lucros exorbitantes. Utilizando grandes sistemas de recomendação de dados e algoritmos de aprendizagem de máquinas, são capazes de prever com precisão o comportamento futuro dos utilizadores. Outros aspectos para além da precisão devem ser considerados para o sucesso de tais sistemas. Do ponto de vista da empresa, a Netflix tem vindo a envolver-se com sucesso no capitalismo de vigilância em vídeo streaming. Utilizando o modelo de aceitação de tecnologia para adoptar um utilizador-perspectivo, este documento examina a utilidade dos utilizadores do sistema de recomendações da Netflix. Os efeitos da transparência e curadoria como características da plataforma Netflix, juntamente com o nível de confiança dos utilizadores no sistema, são examinados quanto ao seu impacto na percepção da utilidade e facilidade de utilização do sistema de recomendação, a fim de determinar a intenção comportamental do utilizador de utilizar o sistema e a utilização real do sistema. São explorados os efeitos potenciais da interacção do utilizador como uma potencial característica futura. Utilizando modelos de equações estruturais sobre dados recolhidos dos inquiridos, o documento conclui que a cura e a confiança têm de facto impacto na intenção comportamental e na utilização, enquanto a transparência não tem impacto nos factores percebidos. A interacção dos utilizadores não melhora significativamente a utilidade dos utilizadores. O resultado sugere que o Netflix deveria concentrar-se nas características de cura e de construção de confiança como características diferenciadoras da sua plataforma para sustentar a vantagem competitiva.

Table of contents

Acknowledgements

The completion of this thesis would not have been possible if it were not for a number of extraordinary people. This is for you.

First of all, I would like to express my gratitude to my supervisor, Prof. Boris Durisin for supporting me throughout the process. With his sharp analyses and precise feedback he was able to weed through my at times disorganised thoughts and comments in the shortest amounts of time during our tutoring sessions and provide guidance through logic and structure. At the same time he awarded me a great amount of freedom in discovering my topic and developing a research approach, while being a constant inspiration with his engaging teaching style.

I would also like to thank Professors Miguel Godinho de Matos, Christian Peukert, Raffaele Conti, and Ilídio Barreto of Católica Lisbon School of Business & Economics for sparking my enthusiasm for approaching management, technology and digitalisation with scientific rigour.

Furthermore, a great deal of recognition of my dear friends – old and new – is in order. They continue to inspire me through their kindness, intelligence, humour, and diligence. I am truly lucky to have forged such strong bonds with a diverse group of people so that regardless where I go I will feel at home and have people stand by me, no matter if I have known them for 17 years or 10 months.

I wish to express my deep gratitude to my girlfriend Anna who not only has endured the hardship of distance for me to be able to complete my studies, but has continuously been there for me with her love, encouragement, advice, her emotional intelligence, and an open ear for all my complaints.

Finally, I am eternally grateful to my family. To my father who has not only encouraged me to learn about the world, but also financially enabled me to study abroad and has always been an avid reader of my academic writings. To my big brothers Andras and Peter for being my biggest supporters ever since I was a little child, always providing valuable advice, and inspiring me to travel and study in foreign countries. To my sisters-in-law who cheer me up with pictures of my nieces and nephews when I am feeling down. And to my mother, aunt, and grandmothers who have raised me with all the love and affection one can only hope for.

Thank you | merci | obrigado | grazie | dankeschön | kiitos | köszönöm

List of figures

List of tables

List of abbreviations

A	Attitude towards using
BI	Behavioural intention to use
CAGR	Compound annual growth rate
CB	Content-based
CF	Collaborative filtering
CS	Cold start
CFI	Comparative fit index
Df	Degrees of freedom
EBITDA	Earnings before interest, taxes, depreciation, and amortization
EDT	Expectancy disconfirmation theory
EOU	(Perceived) ease of use
IMDb	Internet Movie Database
H	Hypothesis
IS	Information search
KBS	Knowledge-based systems
KPI	Key performance indicator
OLS	Ordinary least squares
RMSEA	Root mean square error of approximation
RQ	Research question
RS	Recommender system
TAM	Technology acceptance model
TLI	Tucker–Lewis index
TPB	Theory of planned behaviour
TRA	Theory of reasoned action
SEM	Structural equation modelling
SRMR	Standardized Root mean square residual
U	(perceived) usefulness
USP	Unique selling proposition

1. Introduction

In 2006 the mathematician Clive Humby allegedly coined the phrase "data is the new oil" (Flender 2019). The same year the then-DVD rental company Netflix announced a USD 1 m prize for a video recommendation algorithm that would be able to outperform the company's own. In 2007 the company debuted its online video streaming service and has since then transformed its main business model from DVD rental to on-demand streaming (Iqbal 2021). Since 2006 Humby's phrase has trickled down from technology-focused media outlets describing the profit opportunities of data (Toonders 2014) to mainstream media discussing regulation ("The world's most valuable resource" 2017)[1] to policy briefings of the European institutions (Szczepański 2020). According to the Economist (2017) Google's parent company Alphabet, Amazon, Apple, Facebook, and Microsoft reaped profits of USD 25 bn. in the first quarter of 2017 alone. Soshana Zuboff named these business models with data at its foundations surveillance capitalism. Initially developed by Google and subsequently pursued by Facebook, Amazon, and also Netflix it involves acquiring behavioural surplus (data generated as by-product of human-system interaction), the creation of predictive products, and the implementation of behavioural modification. Using machine learning, consumption behaviours are forecasted (Cuéllar & Huq 2020). Technology companies are feeding the collected data into probabilistic models and algorithms to predict future thoughts, feelings and acts of users, and recommend products and services. Recommender systems (RS) modify (or at least influence) the behaviour of users by presenting them with certain links and websites rather than others (Google), filtering content and advertisements on their feed (Facebook), suggesting some products over others (Amazon), or by recommending which movie one would enjoy watching (Netflix). Academic research has mostly considered recommender systems through an information systems perspective, either by classifying the RS according to the algorithms used or on the basis of the recommendation (user or recommended product/service), such as in Hand et al. (2001). Other technical papers have focused on accuracy of recommender systems as primary success metric (e.g., Herlocker et al. 2004; Adomavicius & Tuzhilin 2005).

Scholars such as Gorgoglione et al. (2019) have called for research that puts business-performance as the outcome of RS performance at the centre. RS are after all merely tools in

[1] The cover story of the May 6th 2017 edition of the Economist (which does not state the names of its authors for individual articles) was *The world's most valuable resource is no longer oil, but data*, showing a number of off-shore oil drilling platforms with the logos of large technology companies mounted on them.

the overall busines model of surveillance capitalist companies. A number of authors (e.g., Pu & Chen 2010; Knijnenburg et al. 2012; Pu et al. 2012; Cremonesi et al. 2013) suggest to consider user-centric key-performance indicators (KPIs) that should be inherent in RS features. As users are becoming increasingly aware that companies are collecting data on them with or without their explicit consent, trust in systems develops into a central issue and source of customer loyalty (towards trustworthy systems). System characteristics such as transparency and curation are aimed at building trust and user friendliness.

As many technology companies, Netflix claims to be user-centric. The data of the customer speaks for itself and that it does so more objectively than the opinion of a manager or CEO. Having entered the on-demand streaming business in 2007, Netflix quickly established personalised recommendations as its USP. At the beginning of 2021, Netflix had approximately 200 m users worldwide (Amazon Prime Video coming second with 150 m) and accounted for 34% of streaming minutes in the US (the country with the most Netflix users). In 2013 Netflix yet again changed its strategic course of action by placing an increased emphasis on original content for the coming years. Netflix originals such as *House of Cards*, *Orange is the New Black*, *Stranger Things*, *The Crown* or *Ozark* have been celebrated as successes, both commercially and in terms of critical acclaim (Iqbal 2021). Additionally, these in-house productions are more expensive than traditional series produced for TV, with *The Crown* leading the ranking of Netflix's most expensive shows. At the same time the competition is seeking to gain a foothold in the market. Streaming services such as Amazon Prime Video, Disney+ and Apple TV+ are trying to differentiate based on pricing, technological features or their own original content (Seale 2019). Netflix's investors demand appropriate returns for such capital expenditures. According to Lopez et al. (2010), when applying RS commercially correctly (that is in a user-centric approach) the utility of users should go up and usage should increase. This should lead to positive corporate financial KPIs.

However, given the strategic shift from personalisation to original content and increased investment into the latter, naturally the question arises whether the personalisation RS of Netflix still in fact creates customer utility and whether that is translated into added value for the company. Thus, the purpose of this paper is to investigate the functionality of the Netflix RS from a user-centric perspective. To do that it utilises the technology acceptance model (TAM) proposed by Davis et al. (1986). In doing so it does not only examine whether the RS creates utility for users and the company, but also sheds light on what features of the Netflix RS induce acceptance of the technology, leading to utility. Furthermore, in an approach to anticipate future strategic choices of the company given the features of competing RS, the paper attempts to identify if adding user interaction as new feature leads to increased utility. Hence, the following research questions are formulated:

> *RQ1: Does the acceptance of the Netflix recommender system lead to increased utility for customers in the form of usage and for the company in the form of revenues and profits?*
>
> *RQ2: What features of the Netflix recommender system are the determinants of customer utility?*
>
> *RQ3: Can customer utility be increased by extending the Netflix recommender system with user interaction features?*

The findings of the paper indicate that in fact the Netflix RS creates customer utility in the form of increased usage and translates into financial success for the company and its shareholders. Trust and curation are found to be determinants of customer utility while transparency features partially lack the full impact that Netflix intends. User interaction was not found to improve customer utility.

The paper adds to the existing literature on recommender systems by adopting a user-centric view and applying the technology acceptance model to the specific case of Netflix, a company previously at the forefront of generating customer utility and company value through recommender algorithms. On an academic level, the research advances the understanding of what features of recommender technology lead to technology acceptance and customer utility and whether a strategic shift affects customer utility through perceiving system characteristics differently. Through the analysis of a specific company, the findings of the paper provide insights for both Netflix and other streaming providers how to leverage technology acceptance of recommender systems to generate sustained business value and what strategic directions to follow in the future under consideration of the competitive dynamics.

The paper is structured as follows: Section 2 reviews the relevant literature on recommender systems and the technology acceptance model. Section 3 outlines the conceptual model, the constructs, and hypothesised relationships along with the survey design for collecting sample data. Section 4 details the data analysis including relevant testing of the measurement model and structural equation modelling of the hypothesised relationships in the structural model. Section 5 provides concluding remarks by revisiting the research questions, addressing the limitations of the paper, and offering suggestions for possible extensions.

2. Literature review

2.1 Recommender systems

The following section reviews the academic literature on recommender systems. A substantial part of the research originates from the information systems sphere, given the algorithmic nature of RS. Nevertheless, it is important to highlight their key technological features and their central statistical measure of quality, accuracy. The paper also sheds light on alternative quality measures proposed by authors. These more user-centric measures are also found to be central elements of the specific Netflix RS, as Netflix prides itself to be a user-centric company. Hence, the technological specifics of the Netflix RS, are suggested to be determinants of customer utility and usage.x

2.1.1 Taxonomy and definitions

According to Resnick and Varian (1997 p.56) “recommendation systems assist and augment the natural social process” of seeking support to make a decision, e.g., when choosing a book based on the suggestions of a professional or a friend. Schafer et al. (2001) note the evolvement of RS as a consequence to the frustration of users created by information and choice overload caused by the increased production of data. A number of scholars have provided definitions for RS. Stohr and Viswanathan (1998) suggest that RS provide decision support by offering insights about relative payoffs of different courses of action. Burke (2002) describes the purpose of RS as personalised guidance for users to find interesting or useful objects among a wide range of options. Meteren and van Someren (2000) touch upon the algorithmic nature of RS, identifying the process as classification task. The authors distinguish between explicit systems, where users have to express or rate their choices explicitly to receive future recommendations, and implicit systems, that monitor user actions and infer preferences (ibid). The need for users to actively engage with the system may reduce its ease of use, while reluctancy to provide information leads to diminished accuracy (Liang et al. 2007). RS can further be classified according to the basis of the recommendation as proposed by Hand et al. (2001) into non-personalised recommendations (NP), collaborative filtering (CF), content-based (CB) RS or hybrid RS. NP RS provide the same recommendations to every user, based on average ratings or listed reviews of products (Poriya et al. 2014). CB RS suggest items similar to ones used before by the user, by creating features and attributes and matching them to user profiles, recommending the closest matches. CB RS face the issue of overspecialization by recommending items similar to what the user had already enjoyed, thus lacking the possibility of recommending serendipitous finds (Schafer et al. 2007). In contrast, CF collects large quantities of behavioural data from different users and predicts preferences of new items based on the relationship between users and items grounded in feedback implicitly or explicitly given on previous items by similar users. CF suffers from cold start (CS) for new users and new items, due to the large number of historical data points required to base the recommendation on which new users and items lack. Hybrid models combine all approaches (Wei et al. 2017). Sridevi et al. (2016) survey previous literature on RS and provide a classification that goes beyond that of Hand et al. (2001). The authors find literature on multi-criteria RS (incorporating other preferences than purely accuracy, as the next section shows) and on knowledge-based systems (KBS). KBS provide complex recommendations for changing user preferences depending on time and context, noting the necessity of feedback loops between users and system (Sridevi et al. 2016).

2.1.2 Quality measures of recommender systems

Different authors (Gorgoglione et al. 2019; Pu et al. 2012) have noted that literature on RS is largely focused on machine-learning performance (Herlocker et al. 2004; Adomavicius & Tuzhilin 2005), measured by accuracy. While accuracy is the primary performance metric of RS and is suggested to lead to customer value creation and loyalty (Lopez et al. 2010), other have argued that maximizing this metric alone is not enough to satisfy diverse user demands, or generate sustained product sales (McNee et al. 2006). Gorgoglione et al. (2019) reinforce the need for business-performance-centric research, as ultimately RS are means for companies to provide added value to customers and thus generate competitive advantage through differentiation. A number of studies have proposed to complement or even replace accuracy as a performance metric in RS with qualitative measures, such as diversity (Zhang et al. 2012, Soares & Viana 2017, Li et al. 2020), serendipity (Maccatrozzo 2012, Zhang et al. 2012), novelty (Zhang et al. 2012), and customer satisfaction (Li et al. 2020). Maccatrozzo (2012) notes issues of overspecialisation as a way to avoid choice overload, proposing serendipity as an alternative performance measure. Zhang et al. (2012) present a RS for music streaming incorporating diversity, serendipity, and novelty at the expense of reduced accuracy. Kunaver and Požrl (2017) provide a survey on existing literature of diversity in RS and underline the lack of consensus and appropriate measures for diversity, suggesting that often an objective definition is missing and is instead replaced by a subjective understanding of users. Li et al.

(2020) propose adding diversity as a driver of customer satisfaction based on the expectancy disconfirmation theory (EDT) by Oliver (1977), measuring expectations against actual service quality. They suggest that high accuracy will lead to high customer satisfaction through matching expectations and actual results, that however over time become repeated and thus decrease customer satisfaction, thereby requiring diversity to increase the probability of intention to use.

2.1.3 User centricity as goal objective of recommender systems

A number of scholars (e.g., Pu & Chen 2010; Knijnenburg et al. 2012; Pu et al. 2012; Cremonesi et al. 2013) advocate for the inclusion of user-centric criteria in the evaluation of RS quality. Pu et al. (2012) survey a series of papers that establish a relationship between transparency, control and privacy and trust in RS, as well as ease of user-system interaction. Pu and Chen (2010) argue for quality measures from the user's perspective, suggesting an ex-post evaluation of systems by looking at behavioural intentions to adopt the system or refer it to friends, depending on how useful users found the system. Sitar-Taut et al. (2012) define perceived trust as well as data security and privacy as key features of RS user-centricity. Pu et al. (2012) suggest an accuracy-ease of use trade-off when trying to maximise intention to use and user utility. The accuracy of the system must be good enough to be worth the cognitive effort (and thus relatively easy to use) in order to suggest high perceived customer utility and thus willingness to use. Knijnenburg et al. (2012) incorporate experience into the technology acceptance model (TAM) framework by Davis (1986) to examine how and why user experience in RS is created. Such personal dispositions are present already in Stohr and Viswanathan (1998). They remark that concerns about trust, validity, privacy, risk, and performance affect the acceptance of RS. The following section examines the adoption of user-centricity in the RS of Netflix.

2.1.4 Applications of recommender systems at Netflix

Former director of research and engineering at Netflix, Xavier Amatriain (2013) published a paper outlining the key features of the Netflix RS, explaining that the filtered content is presented in an order corresponding to the predicted rank. Highest ranked content is predicted to bring users the most enjoyment (utility). Ranking is a two-dimensional function between popularity (users like what other users like but this alone does not represent personalisation) and predicted personal ratings (despite observed fuzziness in user ratings) of content (depicted in Fig. 1). This provides the foundation of Netflix RS, while other algorithmic features have led to a substantial improvement in accuracy compared to non-personalised recommendations purely based on popularity, as depicted in Fig. 2.

Fig. 1: Netflix two-dimensional ranking function based on popularity and predicted ratings (Amatriain 2013)

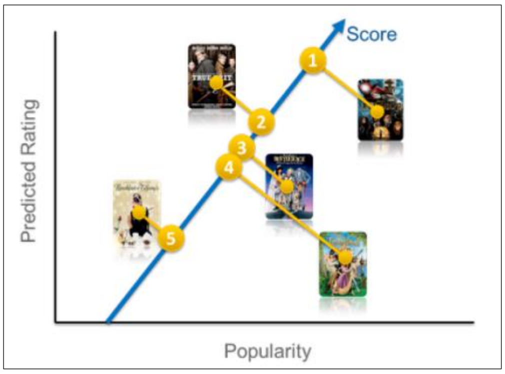
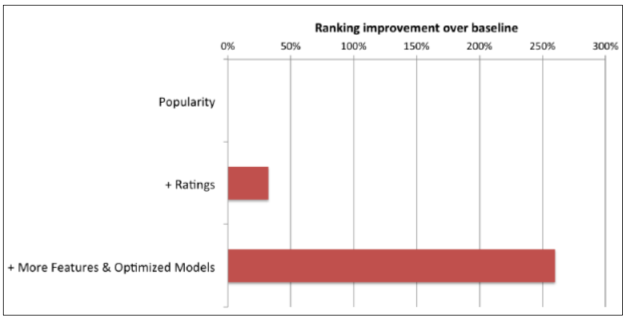


Fig. 2: Performance of Netflix RS when adding features (Amatriain 2013)



Amatriain (2013) notes the importance of transparency for users in explaining why content was recommended, promoting trust in RS as well as encouraging explicit user feedback (e.g. through indicating thumbs up or down on the platform) leading to better recommendations. Conversely, Varela and Kaun (2019) remark, that while most users of Netflix are aware of RS logic (at least to some extent), the advantages of it are expected but not proactively contributed to. Users rarely use the rating feature on the platform but rather rate and review content off-platform (e.g., by posting trailers in social media, rating content on websites such as IMDb, talking to their friends, etc.). This creates dissonance between Netflix and users, as the platform may provide recommendations based on incomplete and biased information and users expectations towards the personalisation features of Netflix are not met. According to Amatriain (2013) Netflix continuously tries to counterbalance this by supplementing user data with 3rd party sources (such as Facebook profile information of users and their friends through using a Facebook account as login credentials on the Netflix platform or enabling the sharing of Netflix content on social media platforms). However, this implicit feedback gathering is not nearly as transparent as the basic recommendation logic explained in Fig. 1, and thus perceived as less trustworthy, demonstrating the need for more explicit feedback gathering, such as through on-platform user-to-user interaction. Gomez-Uribe and Hunt (2015) explain that Netflix uses certain transparent labels for curating the content based on predictions. The *trending now* category proposes content on temporal trends, including seasonally recurring ones such as movies for Valentine's Day or Christmas, and on-off trends, such as natural disasters (where an increased watching of documentaries can be observed). *Because you watched* yields content based on similarity to previously watched videos, a CB-feature. *Continue watching* includes previously unfinished content, ranked by likelihood of resuming. According to Iqbal (2021) curation features are the third most important reason for users to choose Netflix (the other to being related to content and social pressure). The top 10 ranker displays movies that the user is most likely to watch based on the overall prediction of the platform. This feature is not very transparent as users could confuse it with a non-personalised top 10 ranking among users on average. Overall, Netflix attempts to create a certain level of perceived transparency for its users through labelling on the interface. However, users may not fully understand or even misunderstand them. System transparency is not an objective measure but a subjective one that Netflix needs to balance as it may inversely affect the perceived usefulness and ease of use of the RS.

Gorgoglione et al. (2019) observe a number of user-centric characteristics in Netflix's RS. Each row of content recommendations (curated by grouping in different genres) is optimised for relevance and diversity. Additionally, user-experience is optimised based on personalised interfaces (e.g., the labels advertising content adapt based on what mood the user is predicted to be in, such as romantic or thrill-seeking). These user-centric quality measures are aimed at maximising the business objective of user retention, though helping users find engaging content in a very short period of time, increasing usefulness (Gomez-Uribe & Hunt 2015). They further note the correlation of revenue maximisation and maximising customer value. This implies that the Netflix's goal of maximising customer utility also boosts company utility (Maryanchyk 2008). Netflix uses different hybrid RS, depending on if it is a new user one on a free trial or an experienced user. The RS dealing with new and inexperienced users aims to minimize the problem of cold start and provide recommendations as accurate and diverse as possible using ratings from others (CF). In order to do that Netflix also offer a *play something* feature which immediately plays something with a high predicted ranking, without giving any further transparent information on why that title was selected. This feature may be of limited usefulness to users (especially to new users who are more used to following recommendations of real humans and might be sceptical of such systems). However, it is relatively easy to use and the serendipitous selection, if enjoyed by the users, will lead to increased trust in the system. The RS targeting regular users seeks to increase customer loyalty and virality of content by nudging users to recommend new content to other users, or by suggesting content of longer duration, such as series. Depending on the frequency of usage, the RS will recommend more content based on the ratings of others (for less frequent users) or content based on personal usage behaviour. Less frequent users are target for churn-minimising algorithms, while heavy users are subject to systems aimed at revenue maximization through retention and experience strategies (Gorgoglione et al. 2019). Maryanchyk (2008) analyses real Netflix user data and finds that when content is sufficiently reduced and curated by the RS, users revert to average ratings by other users and images of the content shown to make their choice. The paper finds that Netflix RS through transparent display of quality indication leads to customer utility and subsequently reduced churn and increased company profits (ibid).

The theoretical foundations how the features of Netflix truly affect usage and utility will be examined in the subsequent sections, exploring the literature on the technology acceptance model.

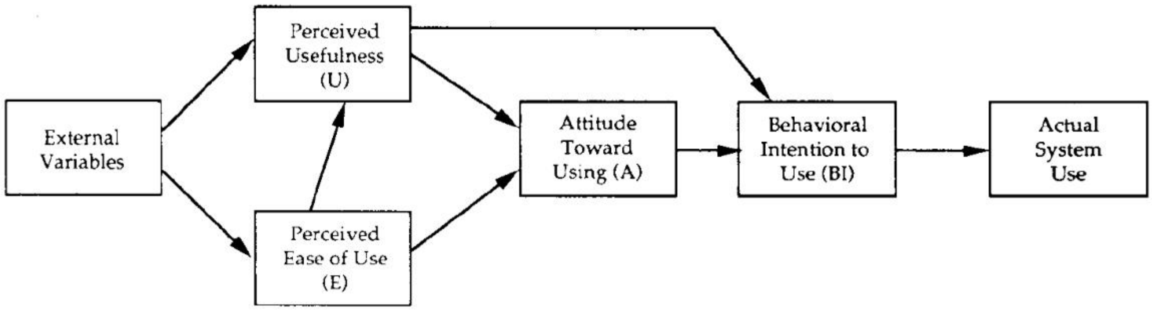
2.2 Technology acceptance model

2.2.1 Theoretical foundations

Davis et al. (1989) argue that new and innovative technologies only improve a product or process relative to the pre-existing status quo (in their case organisational performance), if the technology is actually used. The authors conduct a study that aims to predict acceptance as an outcome of usage intention and usage intention dependent on perceived usefulness, perceived ease of use, and a number of external variables.

Referring to intention models from social psychology, Davis (1986) adapts the theory of reasoned action (TRA) predicting human behaviour (Ajzen & Fishbein 1980), to determine the more specific case of computer usage behaviour, calling it technology acceptance model (TAM, see Fig. 3). TAM demonstrates causal linkages between two key beliefs (perceived usefulness and perceived ease of use), and the user's attitude, intention, and actual use. TAM seeks to not only predict behaviour (forward-looking) but also explain causes (backward-looking) for acceptance or rejection, by observing the impact of external factors on beliefs, attitudes, and intentions. Erasmus et al. (2015) distinguishes between system-specific and user-specific variables.

Fig. 3 Original technology acceptance model (Davis et al. 1989)



Davis (1986) postulates two key beliefs, perceived usefulness (U) and perceived ease of use (EOU), to be determinants of user acceptance of information systems. Davis et al. (1989, p.985) define U as: "the prospective user's subjective probability that using a specific application system will increase his or her (…) performance (…)." . That is, the relation of the performance benefits of usage relative to the effort of using the system (Davis 1989). EOU concerns "the degree to which the prospective user expects the target system to be free of effort" (Davis et al. 1989, p.985). Swanson (1987) notes, that U and EOU are distinct yet inexplicably linked (EOU being a determinant of U). Usage "is determined by behavioural intention to use (BI), but differs in that BI is viewed as being jointly determined by the person's attitude towards using the system (A) and perceived usefulness (U), with relative weights estimated by regression." (Davis et al. 1989, p.985). Lee et al. (2003) find that TAM research has measured BI with usage frequency, amount of time using, and absolute number of usages. The formula to predict BI is:

$$BI = A + U$$

Davis et al. (1989) note that the formula above implies usage intentions if a positive affect is present, (BI→A ceteris paribus). U→BI on the other hand suggests people form intentions about system usage depending on their perception of performance increases, regardless their attitude towards using. The linkage between A and system use has been widely observed (e.g., in Swanson 1982). Usage and usefulness have been linked in previous studies (e.g., Swanson 1987). Lengris et al. (2003) finds a high proportion of studies confirm the stipulated relationship between variables, although with some inconsistencies. Despite clear results suggesting BI as a result of the independent variables of the model, actual use cannot always be predicted. Only seven out of 22 studies surveyed by the authors included both A and BI, while three included only A, and eight only BI. A is determined by U and EOU:

$$A = U + EOU$$

Rooted in Ajzen and Fishbein’s (1980) TRA, TAM assumes that attitude regarding a certain behaviour is determined by relevant beliefs. EOU is assumed to positively impact A, with the effect intending to capture intrinsic motivation caused by EOU (Davis et al. 1989). External variables impact EOU and U, contributing to increased performance and higher efficiency, thus resulting in an effect of EOU on U:

$$EOU = external\ variables$$

$$U = EUO + external\ variables$$

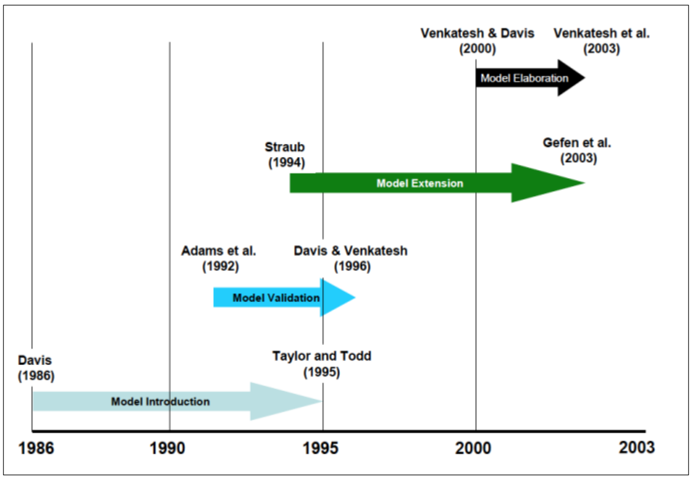
External variables form the linkages between the beliefs, attitudes and intentions postulated by TAM and specific situations (e.g., depending on the technology, industry, voluntariness vs. obligatoriness, and other interventions). With that Davis et al. (1989) provide a framework to study the effect of external variables on user behaviour. Davis (1989) suggests that the importance of usefulness is conceptually sensical, as users are initially inclined to use a system due to the functions it performs for them, and subordinately how simple or difficult it is to get the system to execute certain functions. Davis (1989) further notes, that in terms of causality, EOU is more an antecedent of U than a parallel, direct determinant of usage. Davis finds empirical evidence for the causal relationship of EOU → U → usage (ibid).

2.2.2 Evaluation of the TAM

In a meta-analysis Lengris et al. (2003) find that no study incorporated all the relationships between components (U, EOU, AT, BI, and actual use). TAM is able to explain 40% of system use on average, indicating that significant determinants are missing (ibid). Several papers (Straub et al. 1995; Lengris et al. 2003; Lee et al. 2003) note the dissonance of self-reported use and actually measured use. The authors call for an integration into a wider framework, incorporating human and social change processes. Furthermore, the reviewed research shows that external variables explain system use through mediation by EOU and U. Data on external factors affecting EOU and U provide actionable insights on increasing usage (Lengris et al. 2003).

Lee et al. (2003) identify two main research streams in TAM literature: i). investigating TAM's superiority in comparison to other frameworks and ii). extensions of TAM to other technologies and longitudinal studies (reviewed in the next section) to confirm the model's explanatory predictive ability. A timeline of TAM research is depicted in Fig. 4. Hubona and Cheney (1994) find TAM to be marginally superior to the theory of planned behaviour (TPB), being simpler and more able to explain user acceptance. Taylor and Todd (1995) consider TAM to be less effective than TPB, noting the necessary trade-offs between model intricacy and explanatory power.

Fig. 4: Chronological progress of TAM research (Lee et al. 2003)



2.2.3 Extensions

2.2.3.1 Longitudinal studies

Lee et al. (2003) criticize the cross-sectional nature of TAM for failing to establish causality between variables. Doll and Ahmed (1983) suggest longitudinal studies to analyse the acceptance and usage over time as users become more acquainted with the technology. Karahanna et al. (1999) distinguish between adoption by new users and continuous adoption of pre-existing users. Ventakesh and Davis (2000) confirm the models explanatory power of acceptance over time, measured pre implementation, as well as one and three months post implementation, finding increased BI and decrease usefulness over time, respectively.

2.2.3.2 Different technologies and industries

Adams et al. (1992) apply TAM to five different technologies and find the model to be consistent and valid in explaining acceptance across different systems. Davis (1993) and Subramanian (1994) replicate the results with email and text-editor systems, respectively. Erasmus et al. (2015) examine the acceptance of enterprise resource planning systems. Verma et al. (2018) find system and information quality to be strong determinant of usefulness and ease of use of big data analytics tools in the corporate context. As IT diffused also to consumer applications, so has TAM research. Koufaris (2002) extends TAM to e-commerce environments, incorporating psychological aspects, such as Csikszentmihalyi's flow theory (1975). King and He (2006) compare the effect on BI for office systems vs. e-commerce systems, finding higher effects of EOU on BI for e-commerce and higher impact of usefulness on BI for office applications. Van der Heijden (2004) notes, that the attitude towards using hedonic systems (systems that focus on enjoyment, such as e-commerce or streaming) are more subject to EOU as determinant. A number of studies have addressed technology acceptance in recommender systems (RS). Wang and Benbasat (2005) note the importance of trust in the acceptance of web-based recommendations due to the service-nature of the technology as opposed to the product-focus of previous studies. Hu and Pu (2009) compare acceptance of personalised RS to non-personalised recommendations (e.g., based on average ratings) and find no significant improvements in terms of accuracy. However, users of personalised RS are found to have higher BI to use, mediated through higher usefulness. Pu et al. (2012) examine EOU as antecedent to system acceptance in RS moderated through trust which in return is impacted by transparency and privacy of the system. Pu and Chen (2010) assess the quality of recommendations as determinant of adoption and usage mediated through perceived usefulness amongst other factors. Amertano et al. (2015) test the effect of user skills on RS acceptance, finding stronger impact of usefulness compared to EOU on acceptance. EOU however, increases when users are skilled in using the system in question. Sitar-Taut et al. (2020) contrast RS with simple information search (IS) with regards to attitude towards usage. They find that RS provide higher levels of attitude towards usage than IS as the perceived usefulness is higher in the less time consuming and hence more efficient RS. However, the effect on attitude by increased usefulness is counteracted by privacy concerns of data disclosure that is implicitly or explicitly required to feed the RS to provide high quality recommendations. Hence, the perceived trustworthiness of the RS in handling user data is an antecedent of attitude towards usage (Sitar-Taut et al. 2020).

2.2.3.3 External variables

The evolvement of the TAM also saw the gradual inclusion of different external variables. King and He (2006) distinguish between prior factors influencing usefulness and EOU (such as prior usage or experience), factors from other theories such as trust (e.g., in Jarvenpaa et al. 1998) and contextual factors such as gender or culture (e.g., in Adam et al. 1992; Gefen & Straub 1997). Ventakesh and Morris (2000) examine gender, experience, and subjective norm as determinants to BI. Karahanna and Limayem (2000) find differences in users adoption for the same users of different systems, highlighting system characteristics and individual abilities. Ventakesh and Davis (2000) introduce TAM2, providing antecedents for EOU (such as computer self-efficacy, i.e., the ability to use information technology in general) and usefulness, as well as voluntariness as antecedent to intention to use. Technological diffusion and innovations saw the adaption of TAM for new systems. This required the formulation of new and additional external variables inherent to the specific contexts of these systems, as examined by King and He (2006).

2.2.3.3.1 Trust

Gefen et al. (2003) define trust in the TAM context as the belief in the ability, integrity and benevolence of the trusted party to deliver on its pledges despite the user’s vulnerability and dependence. Reichheld and Schefter (2000) note the crucial importance of trust in e-vendors to attract and more importantly retain users. Hence, increased user trust in the vendor should result in increased intention to use, indicating utility gains for users and vendors alike. Gefen et al. (2003) examine the linkage and find significant impact on BI, suggesting the need for trust-building mechanisms in addition to features increasing EOU and usefulness. Wang and Benbasat (2005) investigate trust as antecedent for the adoption of recommender systems (RS) and find positive moderating effects on intentions to use. They stress the question of benevolence as component of trust as the agency relationship with an online vendor creates uncertainty whether the RS works in favour of the user, the producer, or the vendor itself (ibid).

Wang and Benbasat (2005) note that trust has a bigger impact as determinant of usage intent in settings that involve financial risk (such as e-commerce environments described by Gefen et al. 2003) as opposed to recommendations of RS that do not present an obligation to purchase and consume. Lopez et al. (2010) suggest to add predictability, concerning the expectations towards the object of trust, as fourth dimension to the ones postulated by Gefen et al. (2003). Lopez et al. (2010) stipulate trust in systems as an antecedent to attitude, especially when the use of personal data is involved and users are not fully familiar yet with the system.

2.2.3.3.2 Curation

Kunaver and Požrl (2017) note the impact of diversity of recommendations through curation on output quality and thus usefulness. Kim et al. (2017) suggest that curation of systems increase the ease of use by reducing choice overload. Maccatrozzo (2012) examines the positive effect of curation on usefulness of recommendations, arguing that diversity and serendipitous discovery lead to enjoyment through novel inputs. A system that does not incorporate these factors but rather focuses on the algorithmic-driven aspects of accuracy alone to reduce choice overload, will produce a certain level of ease of use and usefulness. However, such systems are prone to overfitting and do not account for changing moods of users. Thus, according to Maccatrozzo (2012) curation as antecedent to usefulness and EOU incorporates system-specific aspects as well as contextual aspects of the user.

2.2.3.3.3 Transparency

A number of scholars stipulate the positive impact of transparency on system acceptance (e.g., Hayes-Roth & Jacobstein 1994; Gregor & Benbasat 1990). Gregor and Benbasat (1999) find that explanations provided by KBS lead to higher perceived usefulness. Pu and Chen (2010) note the increased purchase intentions of customers for systems with higher transparency. Additionally to direct effects on the perceived usefulness and EOU of systems, system transparency has been suggested to have indirect mediating effects by increasing trust in the system (Pu et al. 2012; Wang & Benbasat 2016) through increased benevolence (as suggested in Gefen et al. 2003). This is achieved by bridging intention gaps. Intention gaps obstruct the formation of trust in KBS as users do not know why a system provides information. Users question the intention of the system and who's utility it is intended to maximise (e.g., the vendor's, the content provider's, or actually the user's), leading to low levels of perceived benevolence (ibid). Transparent (knowledge-based) systems provide reasoning for system-user interaction, thus mitigating users concerns about goal incongruence (Wang & Benbasat 2016).

2.2.3.3.4 Interaction

Sheng and Zolfagharian (2014) integrate user participation into TAM, noting the transition of customer roles from a passive receiver to co-producer. The authors suggest that the time and effort users spend interacting with the system has effects on system acceptance when mediated through usefulness and ease of use. The research finds that user interaction with the system in co-creating to fulfil the system purpose increases usefulness but at the same time makes it more complex, thus reducing ease of use. Hence, the overall effect on technology acceptance depends on the relative impact on usefulness and EOU (ibid). Lopez et al. (2010) stipulate the relevance of user ratings and opinions as well as user to user collaboration for trustworthiness of systems. Kim et al. (2017) note the amplified effects of transparency on system acceptance when supported by networks of users, as other users provide additional sources and validation mechanisms for understanding the reasoning of KBS. Nakka et al. (2020) suggest that the similarity among users is strongly connected to the trust in other user's interactions with the system. Therefore, user interaction is not just a direct antecedent of trust and curation but also indirectly of trust, through increased transparency (by the system indicating how familiar another user is and why).

3. Methodology

3.1 Conceptual model

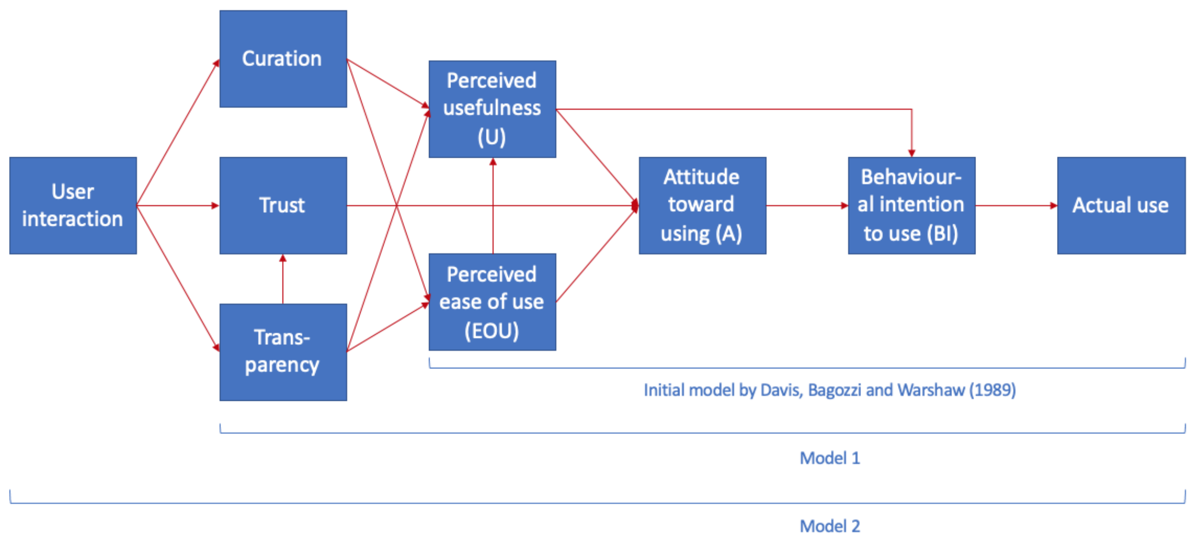
This paper uses TAM as it provides a basis for investigating system use of Netflix and allows to explore the external variables affecting said use, as explained by Davis et al. (1989). TAM was chosen because it has been validated in academic literature covering a number of technological applications over a range of different industries (Burton-Jones & Hubona 2006). Its constructs have been extensively tested for reliability and validity (Chin & Todd 1995; Doll et al. 1999). Meister and Compeau (2002) find that the framework continuously manages to explain 40% of BI and 30% of actual use. BI and actual use serve as indicators of customer utility, as RS add value to customers through improving the decision-making process and providing quality content with the help of personalised recommendations (Lopez et al. 2010; Pu et al. 2010; Amatriain 2013).

The chosen methodological approach adopts the TAM framework by Davis et al. (1989) and extends it with external variables inhibiting individual differences (trust) and system characteristics (transparency and curation) present in Netflix, to observe how these factors affect ease of use, usefulness, and usage (Lee et al. 2003). This is displayed in Model 1. The characteristics of the Netflix platform and user's trust towards the RS are used as antecedents to the key components of the TAM in order to investigate BI and system use and infer customer utility. Additionally, in a second step (Model 2) the model is extended by user interaction to mimic functionalities of the music streaming platform Spotify and test whether that leads to higher perceived usefulness, ease of use, and intentions to use, examining if added features increase added value to customers and thus increase customer utility.

External variables are key to identify the actual usage of Netflix and the utility it thus provides to its users. Lengris et al. (2003) argue that examining external variables are crucial as they drive usage behaviour. Despite their importance, the authors find that only 60% of surveyed studies incorporate external variables (ibid). This paper integrates Trust (see Geffen et al. 2003; Wang & Benbasat 2005), *Curation* (Kim et al. 2017; Kunaver & Požrl 2017), *Transparency* (Wang and Benbasat 2005; Pu & Chen 2010) and subsequently *User interaction* (Sheng & Zolfagharian 2014) as external variables as they provide the basis for Netflix's (respectively Spotify's, and possibly in the future also Netflix's) USP and competitive advantage (Lopez et al. 2010). Demographic indicators such as age (Ventakesh et al. 2003) and gender (Ventakesh et al. 2003), multihoming (Landsman and Stremersch 2011) as well as awareness of the RS

(Lopez et al. 2010) are included to check for moderating effects. Based on TAM, the conceptual model in Fig. 5 is proposed.

Fig. 5: Technology acceptance model of Netflix recommender system, extended model and original model based on Davis et al. (1989)



3.2 Hypotheses

To establish the impact of variables, the paper proposes the following 16 core hypotheses to investigate user's BI and use of Netflix.

3.2.1 External variables

Gefen et al. (2003) and Wang and Benbasat (2005) find trust to be an important antecedent to usage through BI, especially in online settings where system operators can easily partake in opportunistic behaviour. Lopez et al. (2010) find trust a determinant of attitude towards usage. This paper investigates both relationships, given that it has incorporated both BI and A into its conceptual model. Following Lengris et al. (2003) this is rather rare, nevertheless the paper aims at understanding system acceptance and subsequent customer utility along all components of TAM. As trust is regarded as essential component of the proposed model, and A and BI are used interchangeably in studies, measuring the impact on both ensures that no effects will be missed due to selection bias of the author. Trust in a RS should therefore result in a positive attitude and positive behavioural intentions towards using that system.

$H1_a$: Trust has a positive impact on A.

$H1_b$: Trust has a positive impact on BI.

The automatic grouping of titles on Netflix into groups preferred by the user as well as the option to filter titles by different categories should decrease choice overload (Kim et al. 2017) and provide a diverse yet accurate selection of content (Amatriain 2013), thus increasing the probability of system use and thus usefulness (Kunaver and Požrl 2017).

H2: Curation has a positive impact on U.

Curation reduces choice overload and provides different clusters of content that the user can than examine closer depending on their mood (e.g., *action* or *comedy* categories). This serendipitous discovery counters overfitting and makes finding content that the user will enjoy faster and more efficient (Maccatrozzo 2012). Having items on Netflix grouped by genres and other categories on the main page as well as in the categories section should enable users to find suitable titles faster, resulting in increased utility (Maryanchyk 2008).

H3: Curation has a positive impact on EOU.

According to Wang and Benbasat (2005), transparent RS allow users to understand and verify why certain items were recommended, mitigating users' concerns that the RS does not put their interest first when suggesting content. In contrast, an intention gap where users do not understand why a system provides recommendations, may lead to the belief that the RS serves the interest of the provider (the platform or the producer of movies, in the case of two-sided platforms like Netflix). Hence, greater transparency should result in more trust in the RS.

H4: Transparency has a positive impact on trust.

Transparency affects the ability of a RS to convey its inner logic to the user via the user-interface (Pu & Chen 2010). If users understand why a RS offers certain content, they will find the recommendations more helpful, as explanations are openly displayed (e.g., *recommended for you*, *because you watched*, etc.).

H5: Transparency has a positive impact on U.

Similarly, if the reasoning for recommendations is displayed clearly, users might find it easier to operate the system based on recommendations. Without transparent labelling of recommendations and/or their categories, users might be confused why certain content is displayed to them, potentially thinking it is displayed at random or for other (intransparent) reasons (Gregor & Benbasat 1999). Thus, they would find it more difficult to use the system.

H6: Transparency has a positive impact on EOU.

3.2.2 Perceived variables

Referring to the logic of Gefen et al. (2003), this paper assumes that hypothesised paths of TAM as postulated by Davis et al. (1989) also apply to streaming websites in general and the Netflix RS in particular, as users are rational actors when choosing to use a RS. Hence, the more useful and easy to use the RS is, the more utility the users gain from using and the more they will use it.

Improved EOU leads to resources (be it cognitive, human, etc.) being freed up and able to be redeployed elsewhere (Davis et al. 1989), hence increasing the performance of the system as well as of the user in B2C applications. The ease or difficulty of using a system is subordinated to how useful a potential user perceives the system to be (Davis 1989). Thus, finding something faster via the Netflix RS also contributes to a higher probability of using the system and the recommended item.

H7: EOU has a positive impact on U.

A system that is difficult to use will also be met with a negative attitude towards the usage of the system. The system per se can be considered positively, however, the usage by the individual will be regarded negatively. The ease of using a system will positively affect one's attitude towards using the system and reflects the intrinsic motivation of the user towards using caused by EOU (Davis et al. 1989).

H8: EOU has a positive impact on A.

Often positively and usefully perceived outcomes improve the individual's attitude towards those outcomes, as postulated by Davis et al. (1989) and adapted by Gefen et al. (2003) for online environments. High perceived usefulness of Netflix RS should result in positive attitude towards using said RS.

H9: U has a positive impact on A.

3.2.3 Outcome variables

Users of Netflix RS will form behavioural intentions to use such systems if they perceived it as useful, even when not having overly positive attitude towards using said system (Davis et al. 1989). Thus U indirectly affects BI through A (see H11) but also bypasses A and affects BI directly. People will still use Netflix to find a movie if they perceive it as useful in doing so, despite perhaps a negative attitude resulting from previous experience. However, the outcome of BI will depend on the weight of U and A, respectively.

H10: U has a positive impact on BI.

Positive attitude towards usage of systems will result in the intention to use such systems. These variables are so highly correlated that previous researchers have opted to drop one of them due to redundancies and resulting lack of statistical insignificance, as noted by Lengris et al. (2003).

H11: A has a positive impact on BI.

The behavioural intention to use a system (whether to use it again or recommend it to an acquaintance) will lead to actual system use (either by the user themselves or by the acquaintance if being recommended by the system).

H12: BI has a positive impact on actual use.

3.2.4 Model 2

Interaction with other users (that are similar to the user) serve as a trust indicator for the user that others have deemed Netflix RS to be benevolent and its recommendations to be trustworthy. This happens both through a direct impact (Kim et al. 2017) as well as through an indirect effect moderated by increased transparency (Amatriain 2013).

H13: User interaction has a positive impact on trust.

Seeing the curated playlists, favourites and categorisations of other users or friends similar to oneself, induces users to also curate their content by creating lists of favourite content or grouping it by subjective categorisation. This fosters more diversity and serendipitous discovery additionally to the sometimes biased and overfitted recommendations of algorithmic curation (Kim et al. 2017).

H14: User interaction has a positive impact on curation.

Seeing other users' ratings, likes, comments, etc. enhances the transparency of why Netflix RS displays content. Furthermore, it does not only show that content is recommended because user behaviour is similar to that of other like-minded users (Nakka et al. 2020) but also what other users thought of that content. Subsequently, user interaction, through transparency, also affects trust and perceived variables such as usefulness and ease of use (Kim et al. 2017).

H15: User interaction has a positive impact on transparency.

Conditional on the previous three hypotheses, if user interaction has a positive effect on external variables of Model 1, Model 2 is expected to yield a better utility for users, observed through a stronger intention to use.

H16: BI of the user interaction model is higher than BI of the base-line model.

Additionally, the direct and indirect antecedents of BI (EOU, U, A) are tested for difference to the initial model to observe any direct or indirect effects of user interaction. Perceived variables are anticipated to have more positive effects than in the base line model, except for EOU. Following the reasoning of Sheng & Zolfagharian (2014), EOU decreases as interaction grows, making the system more complex and less user-friendly.

$H16_a$: EOU of the user interaction model is lower than EOU of the base-line model.

$H16_b$: U of the user interaction model is higher than U of the base-line model.

$H16_c$: A of the user interaction model is higher than A of the base-line model.

3.3. Survey design

To collect data to be analysed through the proposed conceptual model, a measurement model in form of a survey is created addressing each construct proposed by Davis (1986) and others outlined above. The items measuring constructs are all formulated in a way to address the Netflix RS specifically, in accordance with the methodology of Lopez et al. (2010). However, adapting constructs and items to Netflix from other sources that may or may not investigate the TAM with reference to RS could affect the validity and reliability of the constructs. This matter is examined in section 4.3 on data analysis. Moderator variables *age*, *gender*, and *awareness* are surveyed at the start of the survey. Additionally, the respondents are asked to select the services they use, from a list of video and music streaming platforms popular in Europe, as I anticipate that the majority of responses will come from there. Multiple responses to this question are possible, thus establishing the variable 'multihoming'. Respondents are asked to state how many hours of Netflix they watch per week, in order to measure (self-reported) actual use.

The second part of the questionnaire asks the respondents to recall the Netflix user-interface when searching for/ exploring content. The impact of external variables as well as the components of the original model by Davis et al. (1989) are operationalised through survey items by posting one to three statements concerning each construct, asking the respondents to state their agreement on a 5-scale Likert scale. The items of the survey are based on the empirical work of authors that have written about the TAM and its modern applications such as recommender systems. In accordance with the original model by Davis et al. (1989), constructs of core TAM beliefs as well as constructs relating to external variables are defined and measured corresponding to the characteristics of the underlying behavioural beliefs. Thus, the wording of the operationalising statements inhibits intentions, attitudes, beliefs, and concepts and is adopted from previous empirical literature (detailed in Table 1).

The third part of the questionnaire asks the respondents to imagine Netflix being extended with features of user interaction, such as following each other, following each other's curated lists, seeing what content others are liking and consuming, etc. To help respondents answer this section as accurately as possible without Netflix actually having these features or presenting them with a RS-mock-up (e.g., a minimum viable product in the form of a click dummy), respondents are asked to imagine these improved features to be like the music streaming platform Spotify that incorporates such features. The external variable 'user interaction' as well as the other external variables and constructs of the initial model are measuresd, as above, through statements asking for respondents level of agreement on a 5-scale Likert scale.

The second and third parts of the survey are introduced with screenshots of the Netflix and Spotify user interfaces, providing visual aids and cues about the RS-features of the respective platforms (Fig. 6 and 7).

Fig. 6: Netflix user interface with curation and transparency features, such as popular on Netflix, trending now, because you watched, *and top 10 recommendations (screenshot from the author's personal Netflix account)*

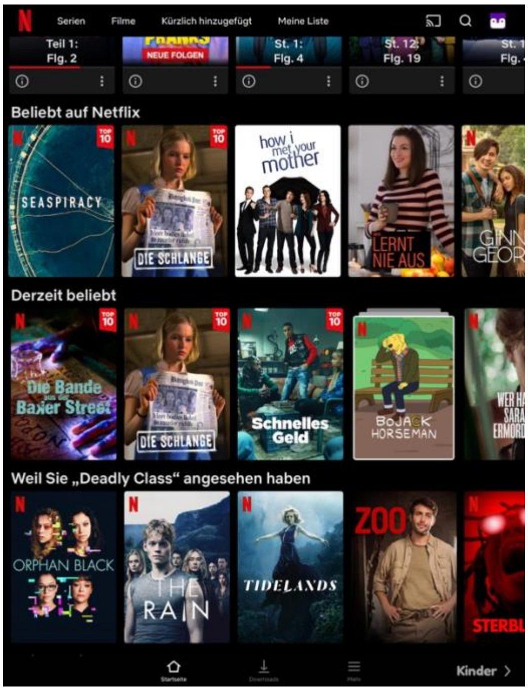
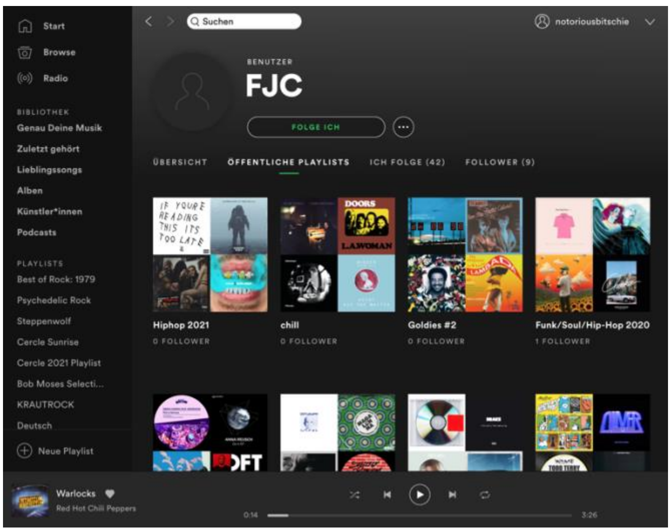


Fig. 7: Spotify user interface with curation and transparency features, such as other user profiles, list of followers of those profiles, and options to like and follow playlists of users (screenshot from the author's personal Spotify account)



Additionally, an attention check was included among the questions posed to respondents towards the final part of the survey, between the statements measuring usefulness. This check asks respondents to respond with "agree" to the statement in order to allow retrospective evaluation if the respondent read the questions and answered them diligently or simply clicked through the questionnaire randomly.

Table 1 below details the measurement model consisting of constructs, items composing the constructs, corresponding questions/ statements for operationalisation of the constructs, and original sources for the questions/ statements.

Table 1: Measurement model with constructs, operationalisations, and original sources

Construct	**Question/ statement**	**Source**
Moderator variables		
Age	Please specify your age	Ventakesh et al. (2003)
Gender	Your gender	Ventakesh et al. (2003)
Awareness	Are you aware that Netflix uses collected data to provide personalised recommendations of content to viewers?	Lopez et al. (2010)
Multihoming	Please select any services that you are using	n.a.
	Model 1	
External variables		
Trust		
Trust 1	Netflix has the ability to understand my needs & preferences about movies/ series	Wang & Benbasat (2005)
Trust 2	Netflix puts my interests first in selecting movies/ series for me	Wang & Benbasat (2005)
Trust 3	Netflix provides unbiased movies/ series recommendations for me	Wang & Benbasat (2005)
Transparency		
Transparency 1	Netflix's reasoning for movies/ series recommendations (e.g., *Netflix original content*, t*op 10 most popular*, *because you watched..* , ratings) is clear to me	Pu & Chen (2010)
Transparency 2	It is apparent to me how Netflix generates recommendations	Pu & Chen (2010)
Transparency 3	I can easily understand Netflix's reasoning process	Pu & Chen (2010)

Curation		
Curation 1	The categories by which Netflix provides its content are appealing to me	Kim et al. (2017)
Curation 2	Netflix organises movies/series into useful groupings	Kim et al. (2017)
Curation 3	Netflix groups its offering in a consistent manner	Kim et al. (2017)
Perceived variables		
Perceived ease of use (EOU)		
EOU 1	My interaction with Netflix's recommendations for content is clear and easy to understand	Amertano et al. (2015)
EOU 2	I discover content more efficiently using recommendation features	Wang & Benbasat (2005)
EOU 3	It was easy for me to understand how recommendations work	Amertano et al. (2015)
Perceived usefulness (U)		
U 1	The recommended shows were tailored to my taste	Amertano et al. (2015)
U 2	The recommended movies were as good as those a friend would recommend	Amertano et al. (2015)
U 3	The technology used by Netflix allows me to find content I like faster	Wang & Benbasat (2005)
Attitude towards use (A)		
A 1	I find it interesting to engage with Netflix' recommendation system	Sheng & Zolfagharian (2014)
A 2	Watching content suggested by the recommendation system is enjoyable	Sheng & Zolfagharian (2014)
A 3	Discovering Netflix recommendations is exciting	Sheng & Zolfagharian (2014)
Behavioural intention to use (BI)		
BI 1	I will continue to use/ try out the recommended content in the future	Wang & Benbasat (2005)
BI 2	I will recommend the Netflix recommendation system to my friends	Wang & Benbasat (2005)
Outcome variable		
Usage	Please specify how many hours per week you watch Netflix (please round to full hours)	Davis et al. (1989)

Model 2		
User interaction		
Interaction 1	Interacting with friends on streaming media requires little effort	Sheng & Zolfagharian (2014)
Interaction 2	Managing my profile and participation on streaming platforms requires little effort	Sheng & Zolfagharian (2014)
External variables		
Trust in user interaction		
Trust_i	Seeing my friends' ratings, comments & content signals unbiasedness to me	Wang & Benbasat (2005)
Transparency in user interaction		
Transparency_i	It is clear to me why the platform shows me friends' ratings, comments & content	Pu & Chen (2010)
Curation in user interaction		
Curation_i	I enjoy the grouped content of other users (playlists, etc.)	Kim et al. (2017)
Perceived variables		
Perceived EOU in user interaction		
EOU_i	Additional collaboration opportunities make streaming recommender systems easy to use	Amertano et al. (2015)
Perceived usefulness in user interaction		
U_i	The recommended movies are those a friend would recommend	Wang & Benbasat (2005)
Attitude towards using in user interaction		
Attitude_i	Using a collaborative recommendation system is enjoyable	Sheng & Zolfagharian (2014)
Outcome variable		
Behavioural intention in user interaction (BI_i)	I will start/ continue to use collaborative features in streaming and recommend it to my friends	Wang & Benbasat (2005)

4. Data analysis

4.1 Data collection and pre-processing

Responses from survey participants were collected over a time frame of two weeks by posting the survey, constructed with Qualtrics software, to a number of social media outlets and web pages. 142 complete responses were gathered additionally to 60 incomplete responses (presumably due to the length of the questionnaire). 9 responses were removed as the participants had failed to pass the attention check question, suggesting that they just clicked through the survey at random, and thus their responses could not be considered valid. A further 6 responses were removed as these respondents had indicated that they do not use Netflix, bringing the final number of valid respondents to 127. The remaining responses were cleaned in Excel of any redundant information generated by the survey tool such as columns for names (this was not asked in the survey for privacy reasons), IP-addresses, coordinates of IP-addresses, etc. Furthermore, the binary variables *Gender*[2] and *Awareness* were recoded from responses returning 1 & 2 to responses showing 1 & 0 (1 standing for *Male* and *Aware*, respectively).

4.2 Descriptive statistics

The subsequent analyses of the data were carried out using the statistical software R Studio. Descriptive statistics yielded interesting initial insights. Users (n=142 displayed in Fig. 8[3]) were asked about the usage of multiple platforms, aimed at identifying multihoming (Cennamo et al. 2018). Most (over 130) indicated that they use Netflix, which is not unsurprising as the survey was advertised under the title *Netflix recommender system survey*. YouTube and Spotify, which offer a freemium[4] business model were ranked second and third, respectively. The content offering of these also appear to be complimentary to each other, Netflix offering movies, series, and documentaries, Spotify music and podcasts, and YouTube diverse free video content uploaded by private and professional users from music videos, to news live streams, to pirated

[2] In the following pages when referring to the variables measured in the models, phrases are capitalised and italicised (e.g., *Gender*). When referring to a theoretical concept, the word is written with small letters and straight font (e.g., usefulness).

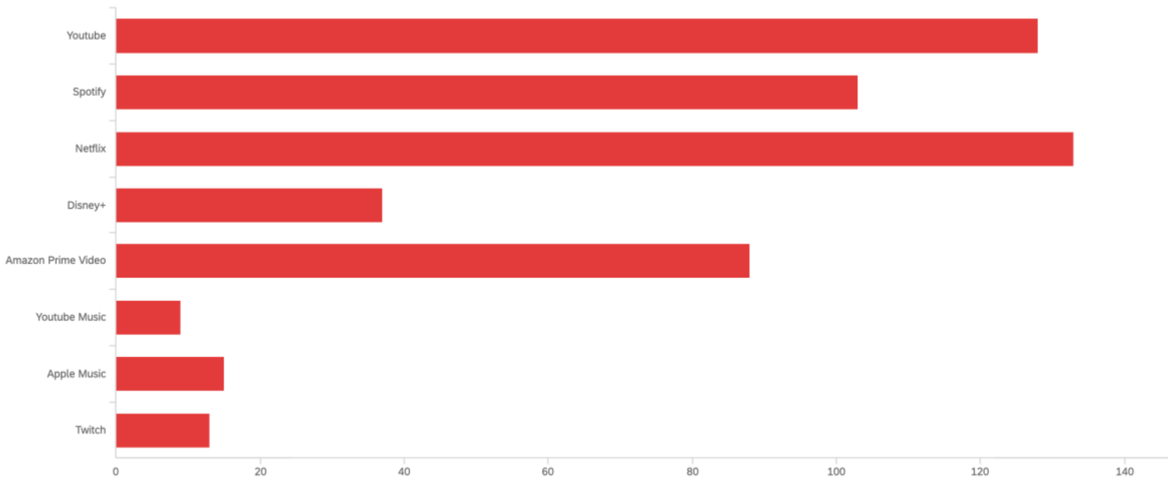
[3] The 142 respondents include those that do not have Netflix or have answered the survey spuriously. Nevertheless, their responses to the multihoming question are valid and thus illustrate that with low or no switching costs users use a number of services that are either complementary or imperfect substitutes.

[4] A freemium business model refers to a free trial or basic version with the continued use or extended features requiring payment (Liu et al. 2015).

content. Only after these three comes Amazon Prime Video, ranked 4th with about 90 users among the participants. Fig. 8 provides details on this.

Multihoming as a variable was ultimately dropped from the analysis as there was too much fuzziness surrounding it. Initially, it was intended to introduce another measure of actual use, namely the relative use of Netflix vis-à-vis other streaming platforms. However, using such a percentage metric as outcome variable could have been problematic due to possible reverse causality. Frequent users of other platforms might perceive Netflix less trustworthy, transparent, curated, useful, or easy to use relative to other similar platforms, such as Amazon Prime, which may lead to churn (both in the short-term during usage by switching and also in the long-term by cancelling a subscription). However, this is beyond the scope of this paper. Such a comparative study could be an extension to this analysis and shall be addressed in the concluding section.

Fig. 8: Number of respondents active on each platform



Summary statistics (see Table 2) provided further insights on the participants. The mean age was 29.4 years. The youngest respondent was 17 years old, the oldest 66, with 50% (n = 71) of respondents being between 25 and 31 years of age.

Gender of respondents was not completely evenly distributed, with 37% of respondents identifying as male and 63% as female. 97.6% of valid respondents reported being aware that Netflix uses systems to recommend content based on data collection and analysis. A mean usage of 7.3 hours of Netflix per week was reported which is a little more than one hour per day, less than the 2019 global average of two hours and the 3.2 hours per day recorded in the US during the covid pandemic in 2020 (Iqbal 2021). The lowest usage was 1 hour per week (respondents were asked to round to full hours when answering this question). The highest usage was reported to be 34 hours, which seems very excessive, given that the 25th and the 75th percentile are 4 and 10 hours respectively. Perhaps the respondent had intended to report 3-4 hours. It is certainly an outlier, but as the other responses of this respondent do not seem out of line, the observation is kept in the data. However, this sheds light on the limitation of self-reported usage (vs. objectively observed or measured usage), addressed in later sections.

Table 2: Summary statistics of measured variables

Statistic	N	Mean	St. Dev.	Min.	Pctl. (25)	Pctl. (75)	Max
ID	127	64	36.806	1	32.5	95.5	127
Age	127	29.339	8.172	17	25	31	66
Gender	127	0.37	0.485	0	0	1	1
Awareness	127	0.976	0.152	0	1	1	1
Usage	127	7.248	5.399	1	4	10	34
Trust	127	3.097	0.593	1.333	2.667	3.333	4.667
Transparency	127	3.297	0.882	1	2.667	4	5
Curation	127	3.43	0.76	1.333	3	4	5
EOU	127	3.483	0.707	1.667	3	4	5
U	127	3.168	0.736	1	2.7	3.7	5
A	127	3.265	0.766	1	2.667	4	5
BI	127	3.283	0.788	1	2.75	4	5
Interaction	127	3.37	0.911	1	3	4	5
Trust_i	127	3.008	0.98	1	2	4	5
Transparency_i	127	3.591	0.867	1	3	4	5
Curation_i	127	3.598	1.018	1	3	4	5
EOU_i	127	3.362	0.914	1	3	4	5
U_i	127	3.22	0.967	1	2	4	5
A_i	127	3.614	0.855	1	3	4	5
BI_i	127	3.315	0.973	1	3	4	5

98% of valid respondents indicated that they are aware of the Netflix RS. As a consequence, the variable *Awareness* is dropped from the moderating variables and henceforth not included in the regression models, as the sample is highly skewered and the information gain from this variable would be limited.

4.3 Reliability and validity of measurement model

The analysis of the data was conducted following a two-stage methodology (Gerbing & Anderson 1988). Prior to testing any structural models, the measurement model (survey), developed and detailed in the previous section, was assessed for reliability and validity. As noted by Nunnally (1978), reliability indicates to what magnitude a measurement item is free of random error. Davis (1986) further explains, that it refers to the proportion of variance in the observed score due to the true score. With an increase in random error, reliability decreases, thus creating unreliable measures, inflating the standard errors of estimated means, leading to more frequent type II errors, further biassing correlation and regression coefficients. Cronbach's alpha is used to indicate reliability and internal consistency of the measured constructs, using a minimum alpha of 0.60 as a benchmark (Nunnally & Bernstein 1994). As additional reliability measure, composite reliability is reported, applying a benchmark of minimum 0.70 (ibid).

Validity of constructs, as suggested by Davis (1986 p.75), has been defined in a number of ways but it overall refers to the extent to which a measure's true score corresponds to the conceptual variable that the measure is intended to operationalise (Fishbein & Ajzen 1975). Construct validity thus concerns the level of systematic variance of a score in relation to the target construct. A lack of validity may result in interpreting output data through the theorised variable instead of the actually measure variable, leading to a higher probability of wrong inferences (Davis 1986 p.75). Factor loadings of minimum 0.4 are employed as measures of convergent validity as recommended by Hair et al. (1998).

The means, standard deviations, Cronbach's alphas, composite reliabilities, and factor loadings for the measurement items and constructs are reported in Table 3. Examining the reliability of the constructs, *Trust* and *Usefulness* were found to have alphas below 0.60. Additionally, *Ease of use* and *Usefulness* exhibited a composite reliability of less than 0.70. *Trust* demonstrated a composite reliability of over 15, thus deemed to be faulty. This was originated in the factor loadings where item *Trust 2* was found to have a factor load of above 8. Repeated coding in R resulted in the same results. Regarding the validity of items, *Trust 1* and *Trust 2* demonstrated extremely low factor loads (< 0.1). Additionally, *Usefulness 2* was found to have a borderline factor loading of 0.397.

Despite the low reliability and validity of *Trust* and *Usefulness*, they were not excluded as they represent key pillars of the model, with constructs validated by previous research (Wang & Benbasat 2005; Amertano et a2015). The item *Trust 2* could have been replaced by another item measuring trust. However, that would have required a lengthier data gathering process, validating each item before publishing the survey. Due to the reluctance of participants to answer questionnaires repeatedly, this would have required a high number of participants to allow for pilots of small samples to check the items and constructs.

Reliability and validity checks for the measurement model were run again two times, with only item *Trust 2*, and items *Trust 1-and-3* in the other run, respectively (see Appendix 7.1). Including only *Trust 2* in the model resulted in a composite reliability of 15.898 and a factor loading of 8.121. Both being larger than 1, *Trust 2* does not appear to be a reliable and valid measure. As only one item was included in this version, Cronbach's alpha was not reported.

When only *Trust 1-and-3* are included, Cronbach's alpha was -0.017, Composite reliability was 0.136, and factor loading calculations returned 0.535 and -0.016. Alpha and CR do not meet the suggested benchmarks of 0.6 and 0.7, respectively. Hence, this variant of the model lacks reliability. *Trust 1* reported an acceptable convergent validity via its factor loading, while *Trust 3* remained behind its expectations.

The constructs for Model 2 were not separately measured for two main reasons. Firstly, the statements of the operationalised items represent only slight modifications of the original items in Model 1, validated for reliability and validity in previous literature and in this section. Secondly, constructs of Model 2 (with the exception of *User interaction*, reported in Table 1) were measured with one item only. This was done intentionally, possibly sacrificing reliability and validity of the measurement model to collect enough data to test the structural model (Model 2). While Model 2 is merely an extension of Model 1, deeming the trade-off to be worth the cost compared to the resulting benefits, a more extensive questionnaire would have resulted in less people answering the survey fully, thus consequently reducing sample size, and increasing standard error and generalisability of the structural model.

Table 3: Reliability and validity measurements of Model 1 and 2

Construct	Mean	Standard deviation	Cronbach's alpha	Composite reliability	Factor loading
Trust	3.097	0.593	0.431	15.898	
Trust 1					0.062
Trust 2					8.121
Trust 3					0.013
Transparency	3.297	0.882	0.83	0.837	
Transparency 1					0.647
Transparency 2					0.848
Transparency 3					0.878
Curation	3.430	0.760	0.763	0.772	
Curation 1					0.733
Curation 2					0.864
Curation 3					0.572
Ease of use	3.483	0.707	0.646	0.662	
EOU 1					0.709
EOU 2					0.467
EOU 3					0.699
Usefulness	3.168	0.736	0.564	0.584	
U 1					0.570
U 2					0.397
U 3					0.711
Attitude	3.265	0.766	0.756	0.758	
A 1					0.688
A 2					0.749
A 3					0.706
BI	3.283	0.788	0.65	0.708	
BI1					0.915
BI2					0.540
User interaction	3.370	0.911	0.793	0.798	
UI 1					0.875
UI 2					0.752

4.4 Fit statistics

A number of test were carried out to evaluate model fit with data. Chi-squared test was calculated to indicate the differences between observed and expected covariance matrices. The fit indices are evaluated based on Hu and Bentler (1999). Incremental fit indices include comparative fit index (CFI, >0.9 cut-off) and Tucker Lewis index (TLI, >0.9 cut-off), comparing the hypothesised model with a baseline model (i.e., one with the worst fit; Xia & Yang 2019). For absolute fit indices (standardized) root mean square residual (SRMR, <0.08 cut-off) and root mean square error of approximation (RMSEA, <0.06 cut-off) are applied to assess how for the hypothesised model is from an ideal model (ibid).

The results of goodness of fit analysis are reported in *Table 4*. Values were computed in R Studio. The initial model emerged as the best fitting revealing the lowest Chi-squared, the highest CFI and TFI and the lowest SRMR and RMSEA. Model 1 demonstrated good incremental fit (TLI remaining below the recommended minimum 0.9) and acceptable absolute fit (very good SRMR of 0.054, mediocre RMSEA of 0.09). Model 2 reported excellent incremental (CFI above conservative 0.95, TLI 0.932) and absolute fit (very good SRMR of 0.054 and RMSEA of 0.05).

As Appendix 7.2 shows, fit statistics were also observed for alternate models including only *Trust 2* and *Trust 1-and-3*, respectively. The *Trust 2* alternate model outperformed the base case (Model 1) in terms of Chi-squared, CFI, and TLI, while lagging behind in SRMR and RMSE measures. *Trust 1-and-3* alternate model, reported a similar Chi-squared as the base case (Model 1), a higher CFI and TLI, but lower SRMR and RMSEA.

Table 4: Fit statistics of original model (Davis et al. 1989), Model 1, and Model 2

Model	**Chi-squared**	**df**[5]	**CFI**	**TLI**	**SRMR**	**RMSEA**
Original model	275.232	28	0.999	0.998	0.027	0.015
Model 1	344.613	28	0.910	0.874	0.054	0.090
Model 2	275.232	28	0.952	0.932	0.054	0.048

[5] Degrees of freedom

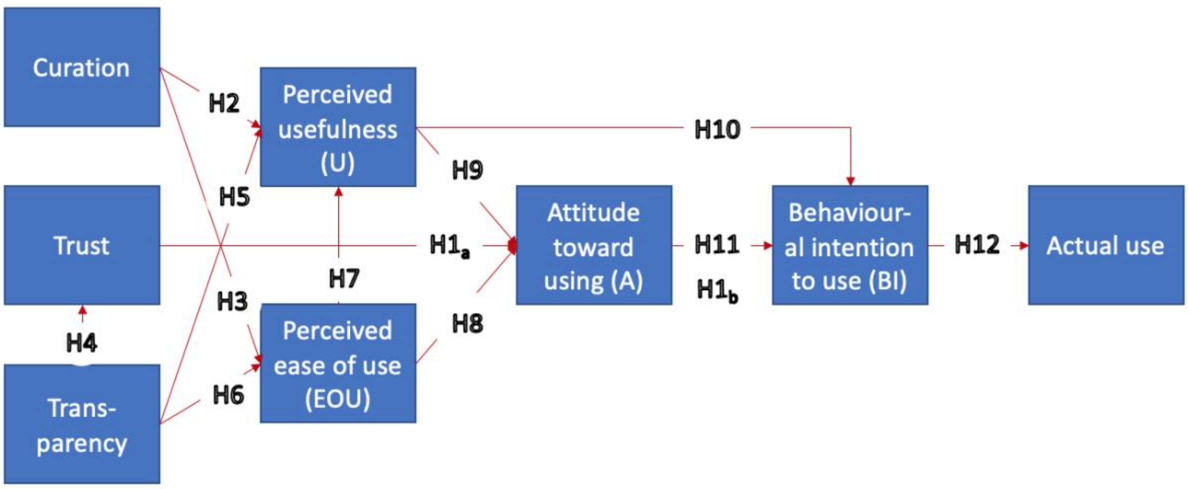
4.5 Structural equation modelling (SEM)

SEM in R Studio was used to evaluate the relationships between items and constructs comprising the structural model. Ordinary least squares (OLS) regression analysis was used to calculate the coefficients of the structural relationships postulated in the conceptual model in the methodology section. The regressions of Model 1 (Netflix RS status quo) and Model 2 (Netflix RS with user interaction) are reported in Tables 4 and 5. Referring to Cohen (1988), a p-value 0.05 was chosen as benchmark for statistical significance (indicated by at least ** in the regression tables). Practical significance of the findings were measured by the coefficients reporting at least 0.30, considered a medium effect by Cohen (1988).

4.5.1 Evaluation of Model 1

Conceptual Model 1 with corresponding hypotheses is displayed in Fig. 9. The path coefficients and their statistical as well as practical significance are reported in the order of the constructs' location in the model, from left to right. The practical and statistical impacts of coefficients are summarized in the OLS regressions (Table 5).

Fig. 9: Hypothesised relationships of Model 1



4.5.1.1 Testing external variables

Trust has a small positive effect on *A* (0.107), not exceeding the threshold for a medium effect. It is not found to be statistically significant. Therefore, $H1_a$ is rejected. The impact of *Trust* was also reported on *BI*, as many studies have used the concept of either attitude or behavioural intention but rarely both (Lengris, et al. 2003), showing both practical (0.346) and high statistical significance. Therefore, $H1_b$ is supported. The same discrepancy is observed in the regression table of Model 2 (Table 6). *Curation* reported a medium effect on *U* (0.301) with high statistical significance. H2 is supported. Similarly, *Curation* was found to have a medium effect on *EOU* (0.361), statistically significant at the 1% level. Therefore, H3 is supported. *Transparency* was found to have a positive effect (0.223) on *Trust*, statistically significant at a 1% level. However, due to the cut-off of practical significance, H4 is not supported. *Transparency* accounted partially (R^2 = 0.11) for the variance in *Trust*. *Transparency* demonstrated a low impact on *Usefulness* (0.104) and is found to be statistically insignificant, thus H5 is rejected. On the other hand, transparency reveals a medium practical significance on *ease of use* (0.319) at a high statistical significance. H6 is supported. Additionally, *Transparency* and *Curation* explained a substantial proportion of the variance of *EOU* (R^2 = 0.48).

Isolating *Trust 2* and *Trust 1 & 3*, respectively, yields slightly different results (see Appendix 7.3). *Trust 2* alone has no significant practical or statistical significance neither on *Attitude towards using* (-0.022) nor *BI* (0.018). *U* and *A* exhibit larger effects on *A* and *BI*, respectively (0.591, 0.406, and 0.509), presumably capturing some of the effects of *Trust*. *U* does no longer have a statistically significant effect on *A*. The explanatory value of the models decreased (R^2 = 0.422 & 0.618 respectively; see Table 10 in Appendix 7.3). The effect of *Transparency* on trust, when only measures through the *Trust 2* item, was found to be statistically not significant and practically less significant (0.116) as the base case reported above (R^2 = 0.013). Including only *Trust 1 & 3* in the OLS regressions produces no significant result for *Attitude* (0.087). The effect on *BI* is statistically significant but practically remains slightly below the cut-off value (0.297). However, this value approximates the statistically and practically significant value of 0.346 of the base case mentioned above and detailed in Table 6. The models using *Trust 1 & 3* are also less able to explain the variance in *A* and *BI* (R^2 = 0.424 and 0.652 respectively; see Table 11 in Appendix 7.3) but better compared to when *Trust 2* is included. Similarly to the *Trust 2* model detailed above, the effects of *U* and *A* on *A* and *BI* are larger, capturing some of the effects of *Trust*, but not as large as the *Trust 2* model. The effect of *U* on *A* remains statistically significant.

4.5.1.2 Testing perceived variables

EOU had a significant practical (0.355) and statistical impact on *usefulness*. *EOU* and *Curation* accounted for 43% of the variance in *Usefulness* ($R^2 = 0.43$). H7 is supported. *EOU* is found to have very little impact on *Attitude towards using* (0.077). That relationship was not found to be significant, thus H8 is rejected. *Usefulness* had a strong and statistically highly significant impact on *A* (0.55). As *EOU* and *Trust* were found to be insignificant, usefulness is able to explain 42.5% of the variance in *Attitude towards using* ($R^2 = 0.425$). Hence, H9 is supported.

4.5.1.3 Testing outcome variables

Usefulness was found to have a statistically significant effect on *Behavioural intention to use*. However, the effect (0.261) was not large enough to show practical significance. Therefore, H10 is rejected. *Attitude* on the other hand was found to have both a significant practical (0.478) and statistical impact. H11 is supported. *Usefulness* and *Attitude towards using* (additionally with *Trust*, as explained above) accounted for 66% of the variance in *BI* ($R^2 = 0.66$). Finally, *BI* was found to have a statistically high impact on *Usage*. *Usage* here is measured in self-reported usage in hours per week. Increasing the behavioural intent by one unit results in 2 hours of additional Netflix consumption per week, equal to an increase of 27% from the mean reported usage of the sample. Therefore, H12 is supported.

4.5.1.4 Effect of moderating variables on Model 1

To moderate the impact on *Ease of use, Usefulness, Attitude towards using, Behavioural intention to use* and *Usage*, *Age* and *Gender* were included in the OLS regressions. The effect on perceived usefulness of male respondents was lower (-0.176) than that of females for *Ease of use* and *Curation*. However, the effect was statistically not significant. Furthermore, men's *Usage* as a result of behavioural intent was almost 1 hour per week lower than women's (-0.856). Also this impact was not found to be statistically significant. Mediating effects of *Age* were very little, with only *Attitude towards using* (-0.015, statistically significant at a 5% level) and *Usage* (-0.109, statistically significant at a 10% level) showing notable results. The concentration of respondents' age between 25 and 31 suggests that with a greater variance in age, more significant effects would have been observable.

Table 5: OLS regression results of Model 1 relationships

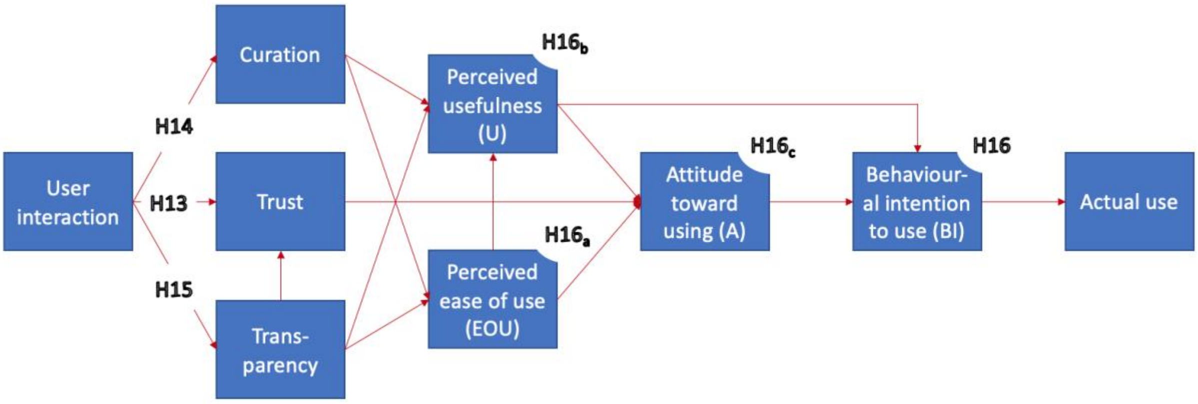
	Trust	EOU	U	A	BI	Usage
A					0.478*** (0.072)	
U				0.550*** (0.099)	0.261*** (0.083)	
EOU			0.355*** (0.092)	0.077 (0.094)		
Transparency	0.223*** (0.057)	0.319*** (0.063)	0.104 (0.07)			
Curation		0.361*** (0.069)	0.301*** (0.078)			
Trust				0.107 (0.115)	0.346*** (0.089)	
BI						2.040** (0.582)
Age				-0.015** (0.007)	0.006 (0.005)	-0.109* (0.056)
Gender				0.002 (0.111)	-0.016 (0.088)	-0.856 (0.945)
Constant				1.349*** (0.396)	-0.336 (0.311)	4.078 (2.671)
Observations	127	127	127	127	127	127
R^2	0.111	0.401	0.433	0.425	0.660	0.134
Adjusted R^2	0.103	0.382	0.409	0.401	0.646	0.113
Residual Std. Error	0.561 (df =125)	0.556 (df=122)	0.556 (df=121)	0.592 (df=121)	0.469 (df=122)	5.084 (df=123)
F Statistic	15.532*** (df =1; 125)	20.446*** (df=4; 122)	18.446*** (df=5; 121)	17.902*** (df=5; 121)	47.014*** (df=5; 121)	6.371*** (df=3; 123)

Note: *p<0.1; **p<0.05; ***p<0.01

4.5.2 Evaluation of Model 2

Conceptual Model 2 with corresponding hypotheses is displayed in Fig. 10. The path coefficients and their statistical as well as practical significance are reported in the order of the constructs' location in the model, from left to right. The practical and statistical impacts of coefficients are summarised in the OLS regressions (Table 6).

Fig. 10: Hypothesised relationships of Model 2



4.5.2.1 Testing user interaction

User interaction was found to have significant practical (0.561) and statistical effects on *Trust_i*. However, it is necessary to remark, that as laid out in the conceptual model, *Transparency_i* also had a mediating effect on *Trust_i* (as in Model 1). The OLS output of R Studio did not display the effect of *Transparency_i* explicitly, suggesting that its effect was absorbed by the variable *User interaction* and/or the constant. *Interaction* (and to some extent *Transparency_i*) explained a substantial part of the variance in *Trust_i* ($R^2 = 0.165$). Therefore, H13 is supported. *User interaction* was found to have a statistically significant effect on *Curation_i*. However, the coefficient (0.276) remains slightly under the threshold. Finally, *User interaction* was found to have little practical effect (0.203) on *Transparency_i*, despite the statistical significance. *Interaction* was not able to explain significant parts of the variances in *Curation_i* and *Transparency_i*, respectively ($R^2 = 0.06$ and $R^2 = 0.046$, respectively). Hence, H13 and H14 are both rejected.

Although not hypothesised explicitly, some impacts of external variables on core beliefs of the TAM in the user interaction model are worth noting. In Model 2, *Curation_i* has a significant statistical and practical effect (0.451) on *EOU_i*. Furthermore, *Curation_i* was found to have negative (although statistically and practically insignificant) effects on *Usefulness_i*. *Transparency_i* had a large and significant impact on *Usefulness_i* (0.404), while *EOU_i* had significant impacts on *Usefulness_i* (0.316) and *A_i* (0.469). As in the previous model, *Trust* was neither practically nor statistically significant as a determinant of *A_i* (0.121) but instead had a strong impact on *BI_i* (0.308). Overall, the independent variables in regressions (6) and

(7) were able to explain 42% and 30% of the variance in *A_i* and *BI_i*, respectively ($R^2 = 0.42$ and $R^2 = 0.30$, respectively).

4.5.2.2 Effect of moderating variables on Model 2

Similarly to Model 1, *Age* and *Gender* were not found to have significant practical or statistical effects when moderating external and perceived variables in the proposed interactive system. Male respondents reported slightly higher perceived *Usefulness* (0.209). *Behavioural intention to use* also had only small effects (0.103). As in Model 1 *Age* was found to have a negative effect on *A_i* (-0.017) and *BI_i* (-0.017). The effects were statistically significant at 5% and 10% levels, respectively. However, the practical effect remains far below the benchmark to be deemed having an at least medium impact. Both effects could have perhaps been more profound with a demographically more balanced sample.

Table 6: OLS regression results of Model 2 relationships

	Trust_i	**Curation_i**	**Trans-parency_i**	**EOU_i**	**U_i**	**A_i**	**BI_i**
Inter-action	0.561*** (0.113)	0.276*** (0.097)	0.203** (0.083)				
U_i						0.078 (0.068)	0.347*** (0.079)
EOU_i					0.316*** (0.1)	0.469*** (0.074)	
Trans-parency_i				0.135 (0.085)	0.404*** (0.095)		
Curation_i				0.451*** (0.073)	-0.022 (0.092)		
Trust_i						0.121* (0.069)	0.308*** (0.085)
Age				-0.005 (0.008)	0.001 (0.009)	-0.017** (0.008)	-0.017* (0.009)
Gender				0.201 (0.138)	0.209 (0.154)	0.094 (0.131)	0.103 (0.163)
Constant	1.724*** (0.345)	2.668*** (0.338)	2.906*** (0.29)	1.337*** (0.436)	0.674 (0.5)	1.890*** (0.36)	1.748*** (0.401)
Obser-vations	127	127	127	127	127	127	127
R^2	0.165	0.061	0.046	0.348	0.295	0.42	0.3
Adj. R^2	0.159	0.054	0.038	0.326	0.265	0.396	0.277
Res. Std. Error	0.843 (df = 125)	0.990 (df= 125)	0.850 (df=125)	0.751 (df= 122)	0.829 (df= 121)	0.664 (df= 121)	0.828 (df= 122)
F Statistic	24.759*** (df=1; 125)	8.129*** (df=1; 125)	5.966** (df=1; 125)	16.252*** (df=4; 122)	10.103*** (df=5; 121)	17.555*** (df=5; 121)	13.060*** (df=4; 122)

Note: *p<0.1; **p<0.05; ***p<0.01

4.5.2.3 Testing statistical differences between the models

To evaluate whether Model 2 truly leads to a higher utility in form of usage, the impacts of perceived variables of Model 1 and 2 are compared. However, actual (self-reported) usage is not measured in Model 2 as it was purely a hypothetical model. The comparison is based on *BI* (as strong antecedent of usage) and the perceived variables of the model by Davis (1986). For that the means of the variables *U_i*, *EOU_i*, *A_i*, and *BI_i* are compared to their counterparts from Model 1. The summary statistics (Table 2) indicate that the reported means for *U_i*, *A_i* and *BI_i* were larger than the corresponding variables of Model 1. Observed *EOU_i* was lower than *EOU*. A one-sided Welch two-sample test is used, to assess if the true means of the variables of Model 2 are larger than the true means of Model 1 variables. The test (results reported in Table 7) returned high p-values, thus not providing sufficient evidence to reject the null hypothesis and means are in fact statistically not different. *BI* in user interactive RS is not higher than in the status quo system and indirect effects of interaction are not impactful enough. Thus, H16 and $H16_{a\text{-}c}$ are not supported.

4.6 Further discussion

The findings of this research provide an application of the TAM to the RS of Netflix. The aim was to observe if the system characteristics of Netflix (transparency and curation) as well as individual differences in the level of trust of Netflix RS lead to user utility in the form of actual usage. Furthermore, it was measured if an extension to Netflix's RS through user interaction would increase perceived antecedents to technology adoption and thus utility. Table 7 below summarises the findings of the analysis with respect to the stated hypotheses.

Trust in the RS does not induce an overwhelmingly positive attitude towards using, contrary to previous studies on trust in TAM (Lopez et al. 2010), that have emphasised the importance of trust in systems handling personal data. In fact, the study by Gefen el al. (2003) observed that vendors have nothing to gain by cheating, thus establishing a basic level of trust. In contrast, the observed means for trust in both models were the lowest among all measured variables (Table 2). Given the past experience of respondents with Netflix, they seem to think that the company behind the RS does have something to gain by not being fully accurate but instead optimising for other parameters, such as revenue maximization, in line with Gomez-Uribe and Hunt (2015). The lack of transparency regarding objectives creates precisely the lack of trust through goal incongruence that Amatriain (2013) claims Netflix tries to avoid through labelling the groupings. The research has shown that trust is not a significant antecedent of attitude towards using, which may be rooted in the users' cognitive inseparability of the RS from the platform. Users may opt to ignore recommendations and only consume content recommended by friends, family, etc. and thus, watch content they find interesting and have a positive attitude towards using it in the future (a sort of confirmation bias). Hence, users may not trust the system but they have a positive attitude towards using (the platform but not specifically the RS). As the recommendations are voluntary and no immediate financial risk is involved in the decision to accept the recommendation (due to the all-inclusive subscription model of Netflix as opposed to Amazon Prime Video for example, where some content requires an additional lump sum payment to rent or purchase on top of the recurring subscription fee), trust is a less impactful determinant of attitude, consistent with Gefen el al. (2003). However, trust was found to have a significant impact on BI, in line with the findings of Wang and Benbasat 2005 and Gefen et al. 2003. This suggests that the medium-term financial risk involved (the monthly or yearly subscription payments) do make users consider whether it is worth to continue using a system (or even recommending it to friends), as noted by Wang and Benbasat (2005). This discrepancy between the effects on attitude and BI may hint at the redundancy of using both constructs in the TAM, as surveyed in other literature by Lengris et al. (2003).

Curation increases both ease of use and usefulness of the RS, confirming the relation suggested by various authors (Maccatrozzo 2012; Kunaver and Požrl 2017; Kim et al. 2017; Gorgoglione et al. 2019). Users perceive higher levels of usefulness of the RS through grouping contents in different categories and thus increased likelihood that following recommendations will result in content to their liking. This happens due to a reduction in information overload and reduction of entropy (Kim et al. 2017), while improving recommendation quality through diversification, solving overfitting (Kunaver & Požrl 2017; Gorgoglione et al. 2019). At the same time splitting the recommendations into categories increases EOU and thus reduces the effort (cognitive but also the time of searching) as it makes choice easier. It is worth noting that it affects the perceived usefulness and ease of use. Netflix RS changes the categories of curation dynamically. Hence, a temporal component may be worth introducing, investigating how respondent's perception of usefulness and EOU changes with changing curation. This lack of persistence is in contrast to Kim et al. (2017) who note its importance for RS effectivity.

Transparency is not a significant antecedent of Netflix RS usefulness, in contrast with findings by Gregor and Benbasat (1990), as the determinants lie elsewhere (EOU and curation for example). Varela and Kaun (2019) provide explanation for this, noting that the feedback the system requires for useful recommendations is instead provided off-platform, discussing content with friends. Thus the system is transparent in providing recommendations (to some extent as explained above in the paragraph concerning trust) but not fully accurate due to missing information, thus lacking usefulness. The perceived subjective transparency of the RS is not enough for users to believe that it will provide recommendations to their taste. On the other hand, transparency has a large impact on the perceived EOU of the system, as the explanation why content is recommended reduces the cognitive effort of the user in having to make assumptions whether they would enjoy the content based on its similarity to previously consumed movies/series, strengthening the RS's roles as decision-support system, in line with Pu and Chen (2010). Transparency as a signalling tool of the RS's competence and integrity (Wang & Benbasat 2005; Wang & Benbasat 2016) does not have a significant *practical* impact on users' level of trust, opposed to what Pu et al. (2012) note. However, the relationship is *statistically* significant. Thus, the transparency feature of Netflix fails its role as trust-provider, despite this explicit aim as explained by Amatriain (2013).

Having established the effects of system characteristics and individual trust levels on perceived factors, the question whether Netflix RS leads to customer utility needs to be addressed by looking at the paths of the core TAM model (H8-12). While not explicitly surveyed, it is assumed that the respondents were repeated/ regular users of Netflix (thus not using a free trial month). Therefore, it can be speculated that they had already established a perception on the ease of use of the RS and the platform in general (as Netflix praises its user friendliness), through the perceived quality of the system and information, as noted by Verma et al. (2018).

EOU has the highest observed mean of all variables measured, indicating the user friendliness of the RS (and the platform as the two were not distinguished in the survey). Its significant effects on usefulness are in line with Davis (1989) and subsequent research meta-analysed by King and He (2006).

On the other hand the perceived likelihood that the RS will return content that the user will enjoy (usefulness) is a strong determinant of attitude, while EOU is not found to be significant, reflecting the correlated and causal relationship of EOU and usefulness, as usefulness captures the effects of EOU on attitude. The lacking significance of EOU's effect on A is in contrast to the findings of Davis et al. (1989), suggesting that system use does not generate sufficient intrinsic motivation to positively and directly impact attitude. Furthermore, the lacking causal relationship is incongruent with the results of van der Heijden (2004) regarding the higher impact of EOU on attitude in hedonic systems. However, the findings are in line with Koufaris (2002).

As attitude is the major influence on the behavioural intent to use, usefulness as a bypass to attitude is not a significant antecedent, contrary to King and He (2006) who note a strong link (which they find because they do not survey models that include attitude). This supports the findings of Lengris et al. (2003) that using both attitude towards usage and BI is redundant. As the effect of attitude on BI and the effect of BI on usage is significant and large, and the model predicting BI explains 66% (see Table 5) of the variance in BI (compared to 40% on average as reported by Lengris et al. 2003), the Netflix RS, through the indirect effects of its system characteristics on perceived variables, is deemed to lead to customer utility.

The added feature of user interaction would increase trust significantly, as users view their peers as trustworthy sources of authority in terms of competence and integrity (Park et al. 1981). This is in line with Tintarev and Masthoff (2007) and Kim et al. (2017). The effect on curation remains behind the expectations, in contrast to Kim et al. (2017), who emphasise the positive amplification of curation through interaction via social networking features and browsing other people's content. The impact is statistically significant but the practical effect (0.276) fell below the benchmark of 0.30 by Cohen (1988). Similarly, the effects of interaction on transparency (and subsequently on trust), as noted by Kim et al. (2017), are not observed. The expected effect of more curation did not have the desired effect on transparency and subsequently trust that user feedback is supposed to encourage, as noted by Amatriain (2013). Instead, it is likely that users interact off-platform already and thus have difficulties perceiving increases in transparency features in an on-platform interaction setting, as users at the moment do not feel the usefulness gains of transparency due to the dissonance observed by Varela and Kaun (2019). Hence, the improvements in the perceived variables were not large enough to be statistically significant and therefore the proposed improvement to Netflix RS customer utility through introducing user interaction is not found to be present.

Though EOU in interactive RS is lower than EOU in standard systems (see Table 2), in line with Sheng and Zolfagharian (2014), also this difference remains statistically insignificant. The means of usefulness, attitude toward usage and BI were all measured to be higher in the interaction model compared to the initial measurement model of the Netflix RS. However, statistical testing of the means has shown that there is not enough evidence to reject the null hypothesis that means are the same. What remains unanswered is whether this was due to people actually not gaining utility from user interaction features on streaming platforms (such as

Spotify, which was presented to the respondents as proxy) or because respondents could not comprehend the suggested features due to a lack of a minimum viable product (e.g., in the form of a website mock-up).

Table 7: Results of hypothesis testing

Hyp.	**Path**	**Coefficient/ mean**	***T*-value**	***P*-value**	**Signi-ficance**	**Support for hyp.**
H1a	Trust → A	0.107	0.935	1.294	n/a	No
H1b	Trust → BI	0.346	3.906	0.0002	0.01	Yes
H2	Curation → U	0.301	3.860	0.0002	0.01	Yes
H3	Curation → EOU	0.361	5.207	0.007	0.01	Yes
H4	Transparency → Trust	0.223	3.941	0.025	0.01	No
H5	Transparency → U	0.104	1.484	0.140	n/a	No
H6	Transparency → EOU	0.319	5.120	0.002	0.01	Yes
H7	EOU → U	0.355	3.855	0.0002	0.01	Yes
H8	EOU → A	0.077	0.821	1.520	n/a	No
H9	U → A	0.550	5.587	0.001	0.01	Yes
H10	U → BI	0.261	3.139	0.106	0.01	No
H11	A → BI	0.478	1.073	3.144	0.01	Yes
H12	BI → Use	2.040[6]	3.503	0.001	0.01	Yes
H13	Interaction → Trust_i	0.561	4.976	0.005	0.01	Yes
H14	Interaction → Curation_i	0.276	2.851	0.254	0.01	No
H15	Interaction → Transparency_i	0.203	2.442	0.216	0.05	No
H16	BI<BI_i	3.283 < 3.315[7]	-0.283	0.611	n/a	No
H16a	EOU>EOU_i	3.483 >3.362[8]	1.177	0.120	n/a	No
H16b	U < U_i	3.168 < 3.220[9]	-0.487	0.687	n/a	No
H16c	A<A_i	3.265 < 3.14[10]	-3.43	0.999	n/a	No

[6] Average hours per week of Netflix usage

[7] Measured means instead of path coefficients reported and compared

[8] Measured means instead of path coefficients reported and compared

[9] Measured means instead of path coefficients reported and compared

[10] Measured means instead of path coefficients reported and compared

5. Concluding remarks

5.1 Summary

This paper has attempted to examine whether users gain utility from using the algorithmic recommender system of the popular video streaming platform Netflix. The technology acceptance model by Davis et al. (1989) was used to predict the behavioural intent to use and self-reported usage. As Gefen et al. (2003) note, the underlying logic of relating usage to utility is that users are rational actors, that will increase they usage if they gain utility from using a system. Perceived usefulness and perceived ease of use are key beliefs in technology acceptance (Davis et al. 1989) and are determined by external variables such as contextual factors and system-specifics. Hence, this paper investigates the effects of system-characteristics that Netflix exhibits (transparency and curation) as well as factors from other theories (trust) on user acceptance and subsequently, utility. Finally, in a forward looking approach the operationalisation of Netflix's strategic direction is attempted and an extension to its existing RS suggested. User interaction is added as antecedent to trust, curation, and transparency to hypothesise if a more interactive system would lead to higher customer utility. The paper has contributed to the existing literature on technology acceptance of recommender systems by drawing on the specific case of a very large and prominent video streaming company, that uses algorithmic RS. Thus, the paper has explored whether the user-centric strategy of the company is truly embodied in the features of technology of its core product, whether these characteristics are significant determinants of core beliefs regarding this technology, and if the beliefs induced by system features actually lead to increased usage and utility for users. The findings are summarised as follows:

> *RQ1: Does the acceptance of the Netflix recommender system lead to increased utility for customers in the form of usage and for the company in the form of revenues and profits?*

Structural equation modelling of hypothesised paths has shown that user acceptance of Netflix RS is significantly influenced by beliefs about usefulness and ease of use. Usefulness has a considerable impact on the attitude towards using while attitude strongly affects behavioural intent to use. The modelled variables in all of these models explain more than 40% (and up to 66%) of the variability in the respective outcome variables. Behavioural intent was found to significantly impact actual (self-reported) usage, with usage increasing by 2 hours per week for every increase in behavioural intent. The findings were reported in Table 5. Hence, the characteristics of Netflix RS lead to increased usage and thus, utility. The findings of the research are in line with the non-financial metrics of the company. Following the logic of Gefen et al. (2003) that users are rational actors that consume more if it gives them higher utility, users are gaining more utility from the Netflix RS, given that already more than 80% of streamed content comes from algorithmic recommendations, as indicated by Gorgoglione et al. (2019). Furthermore, this positive relationship is confirmed by the underlying business success of Netflix. Netflix being a user-centric company utilising customer utility as determinant of company utility has resulted in a revenue CAGR[11] of 26.18%[12] and EBITDA CAGR of 33.12% since launching the streaming service in 2007 (Netflix 2008 & Netflix 2021). Shareholders have valued this approach as well as the share price has increased by 10,618% (CAGR of 43.34%) during the same period[13].

RQ2: What features of the Netflix recommender system are the determinants of customer utility?

The perceived ease of use was found to be a significant determinant of perceived usefulness, while usefulness was a major driver of attitude towards using and, to a smaller extent yet statistically significant, of BI (see Table 5). Papers published by employees of Netflix (Amatriain 2013 ; Gomez-Uribe & Hunt 2015) outline transparency and curation as system-characteristics of Netflix as well as trust as an external factor that system features such as transparency are supposed to induce. The features and the individual perceptions of system trustworthiness are supposed to drive the central goal of Netflix: maximizing enjoyment, or in other words, utility through user centricity. Thus, this paper examined the linkages of external variables trust, transparency, and curation to usage through the beliefs of the TAM.

Trust was found to have significant impact on behavioural intentions (albeit not on attitude, as discussed in section 4.5.1.1 of the analysis). The inconclusive effects of trust hint at people's differing perceptions about the financial risk involved (Wang & Benbasat 2005). On the one hand, the perception of immediate risk of financial loss when opting for a recommendation that does not yield enjoyment (and thus would result in negative utility, due to the financial resources wasted) is not given, as content is all-inclusive (as opposed to Amazon Prime Video).

---

[11] Compound annual growth rate.

[12] While the 2007 figures only include a consolidated revenue amount, the 2020 Netflix annual report distinguishes streaming and DVD revenue (the latter being a mere 1% of the former at around USD 240 m). For the purpose of this analysis, streaming revenues are considered for 2020.

[13] Day after release of annual report (1st of April 2008 and 2021, respectively) used as reference date for share price.

This aspect is captured by the insignificant trust-attitude relationship which is a much more short-term decision. However, for users considering the medium-term financial risk (whether it pays off to pay the monthly or yearly subscription) trust seems to play a role, observed in the impact on BI which reflects a longer-term and repeated usage intention.

Curation was found to be a major driver of utility, as it significantly impacts both usefulness and EOU (see Table 5 for details). Thus, the system-specific characteristic curation is a substantial factors in driving usage and customer enjoyment. Transparency however does not manage to impact usefulness, suggesting that while users find it easier to use the RS through transparent interfaces, the information they receive (recommendations) is not useful, as it does not reflect the true intentions of the system accurately. While the effect of transparency on EOU was significant the effect on usefulness remains behind expectations, in accordance with the dissonance Varela and Kaun (2019) note, due to users not providing the system enough feedback, resulting in sub-par recommendations and low effects of perceived transparency on usefulness. This is also reflected by the insignificant transparency-trust relationship. Transparency was shown to have some effects on trust, but failed to produce an impactful relationship suggested by Amatriain (2003). While it is rational that more openness and transparency results in higher trust, this relationship is not observed in the particular case of Netflix RS, as users might not perceive the system as transparent enough to suggest benevolence of the system which in return would induce trust. However, with the increasing consciousness of users regarding the handling of their personal data by web-based companies, the transparent display where a recommendation is coming from and if it is based on data gathered from linked social media profiles is expected have a larger impact on trust in the future.

Therefore, this paper finds the source of user utility in curation feature characterising the system and in the users' individual perception of trust, depending on the (perceived) financial risk involved. The effects of transparency – contrary to Netflix's objective – does not impact utility to the assumed extent (the effect remaining ambiguous and depending on the individual weights of usefulness and ease of use on BI). It becomes clear that while a transparent interface *should* be an antecedent to usefulness and ease of use, and provide reciprocal feedback to users and the system, thus serving as determinant of increased usage and utility, Netflix fails to establish the desired level of *perceived* transparency. Thus, as indicated in the section on business implications, the company faces the strategic trade-offs, whether to increase transparency, perhaps at the expense of user-friendliness, through having more labels on the interface.

RQ3: Can customer utility be increased by extending the Netflix recommender system with user interaction features?

As some streaming platforms already integrate user to user interaction and emerging technologies such as voice control as well as AR/VR are occupying a larger presence in people's lives, Netflix (the company) is also exploring integrating more interaction into its user experience. To validate whether user to user interaction increases usage and hence customer utility, this paper extends the TAM of Netflix RS by the external variable user interaction. User interaction is found to have a strong impact on trust and no significant practical or statistical effect on transparency, implying that the lacking magnitude of transparency is also carried forward in interaction settings. This hurdle again reflects the issue raised by Varela and Kaun (2019), that people already choose to discuss content off-platform, hence moving this option online adds little perceived value in the transparency domain. On the contrary, this reduces EOU and thus negatively impacts usage and utility. The higher level of trust suggests that users see other users as trustworthy source of authority that substitute (or complement) the medium level of trust in the RS algorithm. While interaction did have an significant impact on Netflix system characteristics and user trust levels, the indirect effects on utility and usage, mediated through the core beliefs, were not significant, suggesting that indeed this (simplified) setting of user interaction does not lead to increased utility.

5.2 Business and managerial implications

A number of managerial and business implications can be derived from these findings. The core concept of Netflix – user centricity - maximises customer success which in turn results in utility for the company and its shareholders through increased revenues, a higher stock price, etc. (Maryanchyk 2008). As Gorgoglione et al. (2019) observe, applying this principle to its product and its recommendation system has resulted in 80% of hours streamed originating from generic recommendations while 20% from recommendations based on searches that did not return the searched item (due to not being available on the platform) and the user choosing a recommended alternative. This has translated into annual revenues of USD 8.7 bn. in 2016 (the year of the recommendation statistic). In 2020 this figure stood at USD 25 bn. (Netflix 2021). As Gorgoglione et al. (2019) rightfully note, the added value of RS to companies is derived not only by prediction accuracy but other user-centric factors. The objectives of the RS ultimately serve to achieve the objectives of the company and not necessarily that of the user. When they diverge, utility is no longer allocated in a win-win setting but rather to the company at the expense of the user. Netflix operates in the increasingly competitive context of the video streaming market in addition to indirect competition from other streaming platforms such as Spotify, attempting to capture the attention of users. Netflix changed its business model and strategy from renting out DVDs to streaming and has seen large financial success and customer loyalty through its personalised recommendations serving as its USP. However, given the market entry of tech majors, such as Amazon and YouTube, as well as innovating movie industry incumbents such as Disney, Netflix must decide on its future strategy. While this paper has adopted a more inward looking view focussing on the resources and capabilities of the company to determine technology acceptance of Netflix's RS, the company needs to understand these finding within the broader scope of competitive dynamics. Gaining competitive advantage through differentiation with original content and locking the user in is increasingly practiced by streaming platforms, the newest potential investment being the planned USD 6.5 bn. acquisition of the MGM movie studio by Amazon to bolster its Prime Video service against the in-house productions of Netflix and Disney (Flint 2021).

In recent years the company has expanded on its original content, increasingly investing into own productions, some of which are highly successful financially as well as critically acclaimed. Netflix produces and airs 1.7 times more original content than the other big streaming platforms combined (Weelright 2020), capturing 61.3% of global demand for original content. These productions represent a substantial investment into what was previously an asset light company. Investors and debtors (Netflix raised USD 15 bn. in debt since 2011) demand to see a positive return on these projects (Iqbal 2021). Hence, it is evident that Netflix increasingly promotes its own shows offline and online, thus raising the question whether algorithmic recommendations still truly reflect a user's interest or even their free will, as they might be nudged to consume (more) Netflix original content, even if the match accuracy or the predicted utility would suggest an alternative choice first. Also as Iqbal (2021) notes, two thirds of users aged 35-54 prefer original content to external movies and series. To ensure sustained user-centricity Netflix must choose whether to continue to base its USP on personalisation alone, shift differentiation fully to original content, or manage to bridge the gap and do both. The latter scenario represents an ambidextrous strategy (Kollmann et al. 2009), where Netflix leverages its core competencies of personalised recommendations with its existing loyal users, while drawing in new users with exiting original content, locking them into the platform through multi-season content, nudging new users to move from a free trial month to a subscription plan, thus increasing revenue. Existing users may opt to try original content through prominent on-platform marketing and transparent labelling as original content and why it is shown. At the same time new users may find enjoyment in discovering diverse content serendipitously through the existing recommendation approaches, thus increasing their utility.

Based on the strategic choices of how to compete, outlined above, Netflix must operationalise its strategy through allocating corporate resources and capabilities (algorithmic prediction vs. content creation and promotion) and adopting the features of its core product accordingly. Given the stipulated importance of transparency for Netflix for both user trust as well as output quality through feedback (Amatriain 2013), the company should aim to increase said transparency if it believes this aspect to be of importance for its systems. However, it is admittedly so that the Netflix RS is a set of increasingly complex algorithms that predict for maximising different objectives. Relaying this transparently to the user would increase trust but most probably decrease usefulness and ease of use due to the information overload, therefore requires a well-balanced approach to explain users why they are seeing certain content to increase trust and usefulness of recommendations. Given the fact that competing streaming platforms such as Amazon Prime Video are more transparent in some aspects (such as displaying a precise rating to the first decimal on a 5-point scale), Netflix must choose its level of transparency as part of a holistic differentiation approach, where the relative payoffs of all features (for example incorporating the effects of financial risk on trust levels of the differing pricing models) are considered, in an analysis evaluated on actual users, e.g., through factor analysis (Child 1990).

Viewed in a larger organisational context, this transparency is not limited to the platform. Netflix should openly promote their original content in marketing campaigns, inviting users to try it despite a potential lack of recommendation fit, thus increasing diversity and subsequently utility. As mentioned in the discussion concerning *RQ2*, the increasing customer awareness of private user data being used without the conscious consent of the user or by being collected from other sources also increases the scepticism and mistrust in IS. Therefore, if Netflix manages to establish trust-building mechanisms (amongst others, through increased transparency), the company stands to gain and keep users, potentially at the expense of competitors. The paper has shown that the competitive advantages of system characteristics of Netflix lead to technology acceptance of the RS, translating into user utility and subsequently company welfare.

If Netflix continues to pursue the strategic path of producing more original content that is specifically made for smart devices such as Smartphones, Smart TVs, and Tables, the company can further operationalise this strategic choice by increasing interaction features. While the paper has not found a significant impact of user interaction on utility for reasons described in previous sections, it is a concept on the rise. AR/VR technologies enable users to interact with content in a more immersing way and Netflix has already experimented with original content such as *Black Mirror: Bandersnatch*, in which the way the story evolves depends on the choices of the user (Damiani 2019). Combining such user to system interaction with user to user interaction enhances the social experience of Netflix and could potentially lead to higher utility for users and more usage in the long run. Potential features such as social watching (groups of people, each person watching on their own device) or reviewing, liking and commenting on friends' movie playlists could lead to more screen time on Netflix. The platform as mediator of user to user interaction (e.g., as visual or audio interface, such as an integrated smart speaker) then takes the role of social actor and should thus increase the trust level in the platform (Wang & Benbasat 2005). According to Porter (1996) such a strategy would reflect a reinforcement of effort, a second-order strategic fit approach, where one system-specific factor (interaction) enhances other factors (such as curation and transparency as well as subsequently recommendation accuracy) and user perceptions (trust). As a further strategic objective the company might consider closer partnerships and integrations of other tech applications such as Facebook or Instagram in order to leverage the users' existing network of friends on these platforms for interaction. This integration of elements of other applications would not only increase interaction and thus user-friendliness, but also help in providing more data, leading to better predictions, both resulting in more utility. The increased amount of user data is then surrendered (more) voluntarily and more transparently compared to Netflix simply acquiring user data on the market, thus increasing trust in the system through the predictability component, as suggested by Lopez et al. (2010). Hence, following Porter (1996) also a third-order strategic fit can be observed through an optimisation of usage, as company activities are geared to increase user friendliness and accuracy through more data collection, both leading to more use, that in return generates more data and attracts more new uses due to network effects and so on. This reflects Porter's view that competitive advantage is derived from a holistic approach to company activities.

5.3 Limitations

Despite the overall coherent results of the paper with respect to previous research both on TAM and RS, as well as literature on and by Netflix, there are several limitations that need to be addressed. The categorisation of limitations follows Chuttur (2009).

5.3.1 Limitations in the methodology for testing the TAM

As noted in other papers (e.g., Lengris et al. 2003) self-reported system use is different from actual system use and might return biased results. Other (self-reported) indicators of usage, such as frequency, were not surveyed. The sample size of participants was too small, considering the recommendation of Kline (1998) to have at least 200, and 10 or even 20 times as many as parameters. The second objective was satisfied as Model 1 (see Fig. 9) incorporated 6 and Model 2 (Fig. 10) 7 independent variables, suggesting 120 to 140 participants. Nevertheless, the actual valid number of participants (n=127) falls short of the recommended minimum. Furthermore, the statements measuring construct items were all phrased positively in order to produce consistency and easy legibility when answering the Likert-scale (*strongly disagree* always being on the left, *strongly agree* on the right). This positive framing without including statements suggesting negative sentiment may have resulted in framing bias. Lastly, survey respondents had to use their knowledge of the Netflix (and Spotify) RS to respond to the best of their ability and rely on the screenshots of the platform interface provided, as no mock-up of an actual RS was provided, which would have reduced some ambiguity in the responses.

While the use of OLS regression analysis in this paper follows the original paper by Davis et al. (1989), a number of authors have used partial least squared (PLS) instead, as it is better suited for small and medium sample sizes and thus non-normally distributed samples (Hair et al. 2011).

5.3.2 Limitations in the variables and relationships present within the TAM

The variables reflecting system characteristics and individual perceptions of Netflix RS were chosen deterministically by the author based on literature on the matter. Only three variables (plus user interaction as the fourth) were included to limit the scope of both the paper and the survey. A larger set of variables inhibiting Netflix's features should have been analysed and prioritised, selecting the most relevant ones. Financial risk as an antecedent to trust could have been observed if more demographic data had been collected, such as household income or employment status. Furthermore, the actual usage of the system is assumed to be dependent on external factors, such as the employment and family situation of the user (presumably a seasoned full-time employee with two kids has less time to watch Netflix than a student). These factors were not measured, possibly biassing the results.

Additionally, no distinction was made between new users and repeated users, despite them being targeted by different algorithms. Reliability and validity checks performed deemed items *Trust 2* and *Usefulness 2* to be compromised. However, they were not excluded for reasons detailed in section 4.3 of the analysis. The inconsistency in trust may have resulted from combining the trust components ability, integrity and benevolence (Gefen et al. 2003) into one composite construct of trust, with one item for each component. Multihoming was not addressed further as it was prone to reverse causality, explained in section 4.2 of the analysis. A more nuanced surveying and analysis of multihoming could have yielded insights on the churn of users to other platforms, its mediating effect on the impact of BI on usage, as well as relative utility derived from Netflix vis-à-vis other services. Responses were demographically homogenous with little dispersion in age. A more diverse sample could have provided insights into age effects in adoption as it is frequently the case with novel technologies (Ventakesh & Morris 2000).

Postulating utility gains as outcome of usefulness and ease of use, via increased BI and usage is somewhat deterministic, despite appearing in literature of both academics (Pu et al. 2012) as well as Netflix employees (Amatriain 2013). The positive relationship of utility and usage is based on the assumptions that users are rational (Gefen et al. 2003). However, this research is limited in the sense that it does not incorporate an objective measure of utility through utility functions or measure relative utility of Netflix RS vs. alternatives that the user might choose to consume. Utility gains might be higher for other systems at a particular moment, but users might opt to stay on Netflix as they have already invested resources (financial, time) to find content (representing a sunk cost).

5.3.3 Limitations in the theoretical foundation for the TAM

In line with Lengris et al. (2003) the inclusion of both attitude and behavioural intentions in the model produced some redundancies and implied some insignificant relationships. Relating to the point made in the previous section, the model assumes actors to be rational, which in reality often is not the case. As such usage might be mediated by a number of external user-specific factors that the model does not incorporate. Furthermore, as noted by Bagozzi (2007), the link between BI and actual use is more spurious than the model suggests, as users still opt out of actual usage after having decided to use Netflix, e.g., if finding appropriate content takes too long (withing 60 to 90 seconds, according to Gomez-Uribe and Hunt 2009).

5.4 Further research

This paper has only touched on some aspects of the acceptance and customer utility of the Netflix RS, leaving much more to be explored by future research. Given the central role of trust and its relations to system-specifics, the research could be conducted with the integrated Trust-TAM model by Wang and Benbasat (2005), examining the effects of system characteristics on usage, mediated by a more diversified and complex trust construct. Furthermore, longitudinal components may be introduced to examine acceptance, and change in utility over time, and gain a better understanding in the utility levels of new vs. seasoned users. Longitudinal studies should also reflect possible strategic shifts of Netflix and investigate the allocation of utility, i.e., whether a shift to a more original content-focused offer increases company utility at the expense of customer utility (e.g., through the goal incongruence of intransparent recommendations) over time.

Additionally, user acceptance of Netflix RS should be explored in the form of comparative studies given the context of market dynamics, where users are simultaneously subscribing and using different platforms (multihoming). As competitors are imperfect substitutes, the negative utility created by one platform through long searches with sub-par results, inaccurate recommendations etc. may create churn to other platforms (Cennamo et al. 2018) and create higher relative or absolute utility. As such, the relationship between BI and usage/utility could be investigated further, taking into account the time lapsed between formulation of intent and actual usage, as well as the switching costs to other systems. This paired with the impact of different business models and hence pricing structures on usage, mediated by the salience of perceived financial risk in trust, provide an interesting basis for a multiple-company case study, embedding TAM as determinant of competitive advantage of systems.

Finally, given the two-sided nature of Netflix and its continued shift to more high profile original content, it is worth investigating the impact of customer utility on the producer side of the platform, namely how the user-centric RS aimed at maximising company objectives affects the long tail of less popular and less recommended content. Hence, relating company, producer and customer utility to direct and indirect network effects provides a multi-perspective advance into the yet little explored research area of technology acceptance of multi-sided platforms.

6. References

Adams, D. A., R.R. Nelson, and P.A. Todd (1992), “Perceived Usefulness, Ease of Use, and Usage of Information Technology: A Replication”, *MIS Quarterly*, 16(2), p. 227-247.

Adomavicius, G. Tuzhilin, A. (2005), "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17(6), p. 734-749.

Ajzen, I., Fishbein, M. (1980), *Understanding attitudes and predicting social behavior*, Prentice-Hall, Englewood Cliffs, NJ.

Amatriain, X. (2013), “Big & personal: Data and models behind Netflix recommendations”, *Netflix*, 1-6.

Bagozzi, R.P. (2007), “The legacy of the technology acceptance model and a proposal for a paradigm shift”, *Journal of the Association of Information Systems*, 8(4), p. 244-254.

Burke. (2002). “Hybrid recommender systems: Survey and experiments”, *User Modeling and User-Adapted Interaction*, 12(4), p. 331-370.

Cennamo, C., Ozalp, H., Kretschmer, T. (2018), “Platform Architecture and Quality Trade-offs of Multihoming Complements”, *Information Systems Research*, 29(2), p. iii-vi, 253-523.

Child, D. (1990), *The essentials of factor analysis*, Cassell Educational, 2nd edition, London, UK.

Chin, W., Todd, P. (1995), “On the Use, Usefulness, and Ease of Use of Structural Equation Modeling in MIS Research: A Note of Caution”, *MIS Quarterly*, 19(2), p. 237-246.

Chuttur M.Y. (2009), "Overview of the Technology Acceptance Model: Origins, Developments and Future Directions”, *Sprouts: Working Papers on Information Systems*, 9(37).

Cohen, J. (1988), *Statistical power analysis for the behavioural sciences*, (rev. edn.), Academic Press, Orlando, FL.

Cremonesi, P., Garzotto, F., Turrin. R., (2013), “User-Centric vs. System-Centric Evaluation of Recommender Systems”, *14th International Conference on Human-Computer Interaction (INTERACT)*, September 2013, Cape Town, South Africa. p.334-351.

Csikszentmihalyi, M. (1975), “Play and intrinsic rewards”, *Humanistic Psychology*, 15, p. 41–63.

Cuéllar, M., Huq, A. Z. (2020), “Economies of Surveillance”, *Harvard Law Review*, 10.02.2020, 133, p. 1280.

Damiani, J. (2019), “Black Mirror: Bandersnatch could become Netflix’s secret marketing weapon. The interactive TV format could give Netflix a needed advantage in a crowded streaming market”, *The Verge*, 02.01.2019, available online at: https://www.theverge.com/2019/1/2/18165182/black-mirror-bandersnatch-netflix-interactive-strategy-marketing [last accessed: 25.05.2021]

Davis, F.D. (1986), “Technology Acceptance Model for Empirically Testing New End-user Information Systems Theory and Results”, Unpublished Doctoral Dissertation, MIT, Cambridge, MA.

Davis, F.D. (1989), “Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology”, *MIS Quarterly*, 13(3), p. 319-340.

Davis, F.D. (1993), “User Acceptance of Information Technology System Characteristics, User

Perceptions and Behavioral Impacts”, *International Journal of Man-Machine Studies*, 38(3), p. 475-487.

Davis, F., Bagozzi, R., Warshaw, P., (1989), “User Acceptance of Computer Technology: A Comparison of Two Theoretical Models”, *Management Scienc*e, 35, p. 982-1003, 10.1287/mnsc.35.8.982.

Doll, W.J., and M.U. Ahmed (1983) “Managing User Expectations”, *Journal of Systems Management*, 34(6), p. 6-11.

Doll, W. J., Hendrickson, A., Deng, X. (1998), “Using Davis's Perceived Usefulness and Ease-of-use Instruments for Decision Making: A Confirmatory and Multigroup Invariance Analysis”, *Decision Sciences*, 29(4), p. 839-869.

Erasmus, E., Rothmann, S., & Van Eeden, C. (2015), “A structural model of technology acceptance”, *SA Journal of Industrial Psychology/SA Tydskrif vir Bedryfsielkunde*, 41(1), Art. 1222, 12 pages.

Fishbein, M. & Ajzen, I. (1975), *Belief, attitude, intention and behaviour: An introduction to theory and research*, Addison-Wesley, Reading, MA.

Flender, S. (2019), “Data is not the new oil”, *Towards data science*, 10.02.2019, available online: https://towardsdatascience.com/data-is-not-the-new-oil-bdb31f61bc2d [last accessed: 29.05.2021].

Flint, J. (2021), “Amazon to Buy MGM, Bagging a Lion to Help Wage Streaming Battle”, *The Wall Street Journal*, 26.05.2021, available online: https://www.wsj.com/articles/amazon-to-buy-hollywood-studio-mgm-for-8-45-billion-with-debt-11622033315 [last accessed: 27.05.2021].

Fornell, C., Larcker, D. F. (1981), “Evaluating structural equation models with unobservable variables and measurement error”, *Journal of Marketing Research*, 18, p. 39–50.

Gefen, D. (2000), “E-Commerce: The Role of Familiarity and Trust”, *Omega*, 28(6), p. 725-737.

Gerbing, D. W., Anderson, J. C. (1988), “An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment”, *Journal of Marketing Research*, *25*(2), p. 186–192.

Gomez-Uribe, C. A., Hunt, N. (2015), “The Netflix recommender system: algorithms, business value, and innovation”, *ACM Trans. Management Information Systems*, 6 (4), p. 1–19.

Gorgoglione, M., Pannielloa, U., Tuzhilin, A. (2019), “Recommendation strategies in personalization applications”, *Information & Management*, 56, 103143.

Gould, J. D., Leweis, C. (1985), “Designing for usability: key principles and what designers think”, *Communications of the ACM*, March 1985.

Gregor, S. and I. Benbasat (1999) “Explanations from Intelligent Systems: Theoretical Foundations and Implications For Practice”, *MIS Quarterly*, (23)4, p. 497-530.

Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1998), *Multivariate Data Analysis*, 4th ed., Prentice-Hall, Englewood Cliffs, NJ.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011), „PLS-SEM: Indeed a silver bullet”, *Journal of Marketing Theory and Practice*, 19(2), p. 139–152.

Hand D., Mannila H., Smyth, P. (2001), *Principles of Data Mining*, MIT Press, Cambridge, MA.

Hauser, J., Shugan, S. (1977), "Efficient Measurement of Consumer Preference Functions: A General Theory for Intensity of Preference", *Discussion Papers 285*, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

Hayes-Roth, F. and N. Jacobstein (1994), "The State of Knowledge-Based Systems", *Communications of the ACM*, 37(4), p. 27-39.

Herlocker, J. L., Konstan, J. A., Terveen, L. G., Riedl, J. T. (2004), "Evaluating Collaborative Filtering Recommender Systems", *ACM Transactions on Information Systems*, 22(1), p. 5–53.

Hsiao, C. H., Yang, C. (2011), "The intellectual development of the technology acceptance model: A co-citation analysis", *International Journal of Information Management*, 31, p. 128–136.

Hu, L.T. and Bentler, P.M. (1999), "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives", *Structural Equation Modeling*, 6 (1), p. 1-55.

Hu, R., Pu, P. (2009), "Acceptance Issues of Personality-based Recommender Systems", *RecSys'09*, October 23–25, 2009, New York, NY.

Hubona, G.S., P. H. Cheney (1994), "System Effectiveness of Knowledge-Based Technology: The Relationship of User Performance and Attitudinal Measures", *Proceedings of the Twenty-seventh Annual Hawaii International Conference on System Sciences (HICSS-27)*, p. 532-541.

Iqbal, M. (2021), "Netflix Revenue and Usage Statistics (2021)", *Business of Apps*, 09.05.2021, available online: https://www.businessofapps.com/data/netflix-statistics/ [last accessed: 25.05.2021]

Jarvenpaa, S. L., Dickson, G.W., DeSanctis, G. (1985), "Methodological Issues in Experimental IS Research Experiences and Recommendations", *MIS Quarterly*, 9(2), p.141-156.

Jarvenpaa, S. L., Knoll, K., Leidner, D. E. (1998), "Is Anybody Out There? Antecedents of Trust in Global Virtual Teams", *Journal of Management InformationSystems*, 14(4), p. 29-64.

Karahanna, E., Straub, D.W., Chervany, N.L. (1999), “Information Technology Adoption across Time: A Cross-sectional Comparison of Pre-adoption and Post-adoption Beliefs”, *MIS Quarterly*, 23(2), p.183-213.

Karahanna, E., Limayem, M. (2000), “E-mail and V-mail usage: Generalizing across Technologies”, *Journal of Organizational Computing and Electronic Commerce*, 10(1), p.49- 66.

Kim, H. M., Ghiasi, B., Spear, M., Laskowski, M, Li, J. (2017), “Online serendipity: The case for curated recommender systems”, *Business Horizons*, 60, p. 613—620.

Kline, R. B. (1998), *Methodology in the social sciences. Principles and practice of structural equation modeling.* Guilford Press, New York, NY.

Kollmann, T. Kuckertz, A., Stöckmann, C. (2009), “Continuous innovation in entrepreneurial companies: exploring the ambidextrous strategy“, *Journal of Enterprising Culture*, 17(3), p. 297-322.

Koufaris, M. (2002), “Applying the Technology Acceptance Model and Flow Theory to Online Consumer Behavior”, *Information Systems Research*, 13(2), p. 205-223.

Kunaver & Požrl (2017), “Diversity in recommender systems – A survey”, *Knowledge-Based Systems*, 123, p. 154–162.

King, W.R., He, J. (2006) “A meta-analysis of the technology acceptance model”, *Information & Management*, 43, p. 4740–755.

Landsman, V., Stremersch, S. (2011), “Multihoming in Two-Sided Markets: An Empirical Inquiry in the Video Game Console Industry”, *Journal of Marketing*, *75*(6), p. 39–54.

Lee, Y., Kozar, K. A., Larsen, K. R.T. (2003), “The Technology Acceptance Model: Past, Present, and Future”, *Communications of the Association for Information Systems*, 12(50).

Legris, P., J. Ingham, and P. Collerette (2003), “Why Do People Use Information Technology? A Critical Review of the Technology Acceptance Model”, *Information & Management*, 40, p.191-204.

Li, Q., Choi, I., Kim, J. K. (2020), “Evaluation of Recommendation System for Sustainable E-Commerce: Accuracy, Diversity and Customer Satisfaction”, *Preprints 2020*, 2020010015 (doi: 10.20944/preprints202001.0015.v1).

Liu, C.Z., Au, Y.A., Choi, H.S. (2014), "Effects of Freemium Strategy in the Mobile App Market: An Empirical Study of Google Play", *Journal of Management Information Systems*, 31(3), p. 326-354.

Liang, T., Lai, H. & Ku, Y. (2007), "Personalized Content Recommendation and User Satisfaction: Theoretical Synthesis and Empirical Findings", *Management Information Systems*, 23(3), p. 45-70.

Maccatrozzo V. (2012), "*Burst the Filter Bubble: Using Semantic Web to Enable Serendipity*", in: *The Semantic Web – ISWC 2012*, Cudré-Mauroux P. et al. (eds), ISWC 2012, Lecture Notes in Computer Science, vol 7650. Springer, Berlin/ Heidelberg, Germany.

Maryanchyk, I., (2008), "Are Ratings Informative Signals? The Analysis of the Netflix Data", *NET Institute Working Paper* No. 08-22.

McNee, S.M., Riedl, J., Konstan, J.A. (2006), "Being accurate is not enough: how accuracy metrics have hurt recommender systems", *CHI Extended Abstracts*, p. 1097–1101.

Meteren, R., van Someren, M. (2000), "Using content-based filtering for recommendation", *MLnet/ECML2000 Workshop*.

Meister, D.B., Compeau, D.R. (2002), "Infusion of innovation adoption: an individual perspective", *Annual conference of the Administrative Sciences Association of Canada (ASAC)*, May 25–28, Winnipeg, Manitoba, p. 23–33.

Nakka, R., Prasad, G., Kumar, R.K. (2020), "Offering Recommendations on Netflix dataset by Associations among Users as Trust Metric", *International Journal of Advanced Science and Technology*, 29(7), p. 989-1000.

Netflix (2008), "Netflix Annual Report 2007", *Netflix*, available online at: https://s22.q4cdn.com/959853165/files/doc_financials/annual_reports/AR_10K_final_2007.pdf [last accessed: 25.05.2021]

Netflix (2021), "Netflix Annual Report 2020", *Netflix*, available online at: https://s22.q4cdn.com/959853165/files/doc_financials/2020/ar/8f311d9b-787d-45db-a6ea-38335ede9d47.pdf [last accessed: 25.05.2021].

Nunnally, J.C. (1978), *Psychometric theory*, 2nd Edition, McGraw-Hill, New York, NY.

Nunnally, J. C. , Bernstein, I. H. (1994), *Psychometric theory,* 3rd Edition, McGraw-Hill, New York, NY.

Oliver, R. L. (1977), “Effect of expectation and disconfirmation on postexposure product evaluations: An alternative interpretation”, *Journal of applied psychology*, 62(4), p. 480.

Park, C. W., Lessig, V. P. (1981), “Familiarity and its impact on consumer decision biases and heuristics”, *Journal of Consumer Research*, 8(2), p. 223–230.

Poriya, A., Bhagat, T., Patel, N., Sharma, R. (2014), “Non-Personalized Recommender Systems and User-based Collaborative Recommender Systems”, *International Journal of Applied Information Systems*, 6(9), p. 22-27.

Porter, M. (1996), „What is strategy?“, *Harvard Business Review,* 74(6). p. 61–78.

Pu, P., Chen, L., (2010). “A user-centric evaluation framework for recommender systems”, *RecSys'11 - Proceedings of the 5th ACM Conference on Recommender Systems*, p. 157-164, 10.1145/2043932.2043962.

Pu, P., Chen, L., Hu, R. (2012). “Evaluating recommender systems from the user’s perspective: survey of the state of the art”. *User Modeling and User-Adapted Interaction*, 22, p. 317-355.

Reichheld, F. F., Schefter, P. (2000), “E-Loyalty: Your Secret Weapon on the Web”, *Harvard Business Review*, 78(4), p. 105-113.

Resnick, P., Varian, H. (1997). “Recommender Systems”, *Communications of the ACM*, 40(3), p. 56-58.

Ricci, F., Rokach, L. Shapira, B., Kantor, P. (2011), *Recommender Systems Handbook*, Springer-Verlag Berlin/ Heidelberg, Germany.

Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S. (2007), “Collaborative filtering recommender systems”, in *The Adaptive Web: Methods and Strategies of Web Personalization*, P. Brusilovsky, A. Kobsa & W. Nejdl (eds.), Vol. Lecture Notes in Computer Science, p. 291-324, Springer-Verlag Berlin/ Heidelberg, Germany.

Schafer, J.B., Konstan, J.A., Riedl, J. (2001). “Recommender systems in e-commerce”, *Data Mining and Knowledge Discovery*, 5, p. 115-153.

Seale, J. (2019), From The Crown to Game of Thrones: what’s the most expensive TV show ever?, The Guardian, 16.11.2019, available online: https://www.theguardian.com/tv-and-radio/2019/nov/16/from-the-crown-to-game-of-thrones-whats-the-most-expensive-tv-show-ever [last accessed: 29.05.2021].

Sheng, X., Zolfagharian, M. (2014), “Consumer participation in online product recommendation services: Augmenting the technology acceptance model”, *Journal of Services Marketing*, 28, p. 460-470.

Sitar-Taut, D., Mican, D., Codruta, M. (2020), “Customer Behavior in the Prior Purchase Stage: Information Search Versus Recommender Systems”, *Economic Computation and Economic Cybernetics Studies and Research October 2020*.

Soares, M, Viana, P. (2017), “The Semantics of Movie Metadata: Enhancing User Profiling for Hybrid Recommendation”, *Advances in Intelligent Systems and Computing*, *WorldCIST 2017*, 1, p. 328-338.

Sridevi, M., Rao, R.R., Rao, M.V. (2016), “A Survey on Recommender System”, *International Journal of Computer Science and Information Security*, Vol. 14, (5).

Stohr, E., Viswanathan, S. (1998), Emergent Structures in the Information Economy, *AMCIS 1998 Proceedings*, 121.

Straub, D. W. (1989), “Validating Instruments in MIS Research”, *MIS Quarterly*, (13)2, p. 147-166.

Straub, D. W., Limayem, M., Karahanna, E. (1995), “Measuring system usage: Implications for IS theory testing”, *Management Science*, 41, p. 1328–1342.

Subramanian, G.H. (1994), “A Replication of Perceived Usefulness and Perceived Ease of Use Measurement”, *Decision Sciences*, 25(5/6), p.863-874.

Swanson, E. B. (1982), “Measuring user attitudes in MIS research: a review”, *Omega*, 10(2), p. 157-165.

Swanson, E. B. (1987), “Information Channel Disposition and use”, *Decision Sciences*, 18(1), p. 131-145.

Szczepański, M. (2020), “Is data the new oil? Competition issues in the digital economy”, *EPRS | European Parliamentary Research Service*, PE 646.117, available online: https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/646117/EPRS_BRI(2020)646117_EN.pdf [last accessed: 29.05.2021].

Taylor, S.,. Todd, P.A. (1995) “Understanding Information Technology Usage: A Test of Competing Models”, *Information Systems Research*, 6(2), p.145-176.

The world's most valuable resource is no longer oil, but data (2017), The Economist, 06.05.2017, available online at: https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data [last accessed: 29.05.2021].

Tintarev, N., Masthoff, J. (2007), *Survey of explanations in recommender systems*, *ICDE Workshops 2007*, p. 801-810.

Toonders, J. (2014), "Data is the new oil of the digital economy", Wired, available online: https://www.wired.com/insights/2014/07/data-new-oil-digital-economy/ [last accessed: 29.05.2021].

Turner, M., Kitchenham, B., Brereton, P., Charters, S., Budgen, D. (2010), "Does the technology acceptance model predict actual use? A systematic literature review", *Information and Software Technology*, 52, p. 463–479.

Van der Heijden, H. (2004). "User acceptance of hedonic information systems", *MIS Quarterly*, 28(4), p. 695–704.

Varela, D., Kaun, A. (2019), "The Netflix Experience: A User-Focused Approach to the Netflix Recommendation Algorithm", in: *Netflix at the Nexus: Content, Practice, and Production in the Age of Streaming Television*, Theo Plothe, Amber M. Buck (ed.), Peter Lang Publishing Group, New York, NY, p. 197-211.

Venkatesh, V., Davis, F. D. (2000), "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies", *Management Science*, 45(2), p. 186-204.

Venkatesh, V., Morris, M. G. (2000), "Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior", *MIS Quarterly*, 24(1), p. 115-139.

Venkatesh, V., Morris, M. G., Davis, G. B., Davis, F. D. (2003), "User acceptance of information technology: Toward a unified view", *MIS Quarterly*, 27(3), p. 425–478.

Verma, S., Bhattacharyya, S.S., Kumar, S. (2018), "An extension of the technology acceptance model in the big data analytics system implementation environment", *Information Processing and Management* 54, p. 791–806

Wei, J., He, J., Chen, K., Zhou, Y., Tang, Z. (2017), "Collaborative filtering and deep learning based recommendation system for cold start items", *Expert Systems With Applications*, 69, p. 29–39.

Wheelright, T. (2020), “Which Streaming Platform has the Most and Best Original Content?”, *Whistleout*, available online at: https://www.whistleout.com/Internet/Guides/streaming-original-content [last accessed: 23.05.2021]

Xia, Y., Yang, Y. (2019), “RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods”, *Behavior Research Methods,* 51, p. 409–428.

Zhang, Y.C., Séaghdha, D., Quercia, D., Jambor, T. (2012), “Auralist: Introducing Serendipity into Music Recommendation”, *WSDM 2012*, p. 13–22.

7. Appendix

7.1 Modified reliability checks for *Trust 2* and *Trust 1 & 3*

Table 8: Reliability and validity measurements of Model 1 with Trust 2

Construct	Mean	Standard deviation	Cronbach's alpha	Composite reliability	Factor Loading
Trust	3.301	0.913	n/a	15.898	
Trust 2					8.121

Table 9: Reliability and validity measurements of Model 1 with Trust 1 & 3

Construct	Mean	Standard deviation	Cronbach's alpha	Composite reliability	Factor Loading
Trust	3.008	0.913	-0.017	0.136	
Trust 1					0.535
Trust 3					-0.016

7.2 Modified fit statistics for Model 1 including *Trust 2* and *Trust 1 & 3*

Table 10: Fit statistics of Model 1 vs. Model 1 with Trust 2 *and* Trust 1 & 3

Model	Chi-squared	df	CFI	TLI	SRMR	RMSEA
Model 1 (initial)	344.613	28	0.910	0.874	0.054	0.090
Model 1 (Trust 2)	288.734	28	0.942	0.919	0.054	0.068
Model 1 (Trust 1 & 3)	344.636	28	0.918	0.886	0.053	0.086

7.3 Evaluation of structural model 1 with *Trust 2* and *Trust 1 & 3*

Table 11: OLS regression results of Model 1 relationships with Trust 2 *(only relevant relationships presented)*

	Trust 2	**A**	**BI**
A			0.509***
			(0.076)
U		0.591	0.406***
		(0.089)	(0.079)
EOU		0.100	
		(0.092)	
Transparency	0.116		
	(0.092)		
Trust 2		-0.022	0.018
		(0.058)	(0.049)
Age		-0.015**	0.004
		(0.007)	(0.006)
Gender		0.002	-0.011
		(0.112)	(0.094)
Constant	2.909***	1.550***	0.153
	(0.313)	(0.402)	(0.329)
Observations	127	127	127
R^2	0.013	0.422	0.618
Adjusted R^2	0.005	0.398	0.602
Res. Std. Error	0.907 (df=125)	0.594 (df=121)	0.497 (df=121)
F Statistic	1.599	17.647***	39.113***
	(df=1; 125)	(df=5; 121)	(df=5; 121)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 12: OLS regression results of Model 1 relationships with Trust 1 & 3 *(only relevant relationships presented)*

	Trust 1 & 3 (composite)	**A**	**BI**
A			0.485*** (0.073)
U		0.561*** (0.096)	0.292*** (0.082)
EOU		0.082 (0.094)	
Transparency	0.224*** (0.057)		
Trust 1 & 3 (composite)		0.087 (0.109)	0.297*** (0.086)
Age		-0.015** (0.007)	0.004 (0.005)
Gender		0.009 (0.112)	0.011 (0.089)
Constant	2.250*** (0.196)	1.382*** (0.391)	-0.234 (0.309)
Observations	127	127	127
R^2	0.109	0.424	0.652
Adjusted R^2	0.101	0.400	0.638
Res. Std. Error	0.568 (df=125)	0.593 (df=121)	0.475 (df=121)
F Statistic	15.226 (df=1; 125)	17.818*** (df=5; 121)	45.338*** (df=5; 121)

Note: *p<0.1; **p<0.05; ***p<0.01