Journal of Promotional Communications

Publication details, including instructions for authors and subscription information: http://

You are more than what you click: Exploring the Impact of Movie Recommender Systems on Consumer Well-being

Marjan Biavaz and Janice Denegri-Knott

To cite this article: Biavaz, M., and Denegri-Knott, J. 2024. You are more than what you click: exploring the impact of movie recommender Systems on consumer well-being. *Journal of Promotional Communications,* 11 (1), 23-52.

Marjan Biavaz

You are More Than What you Click: Exploring the Impact of Movie Recommender Systems on Consumer Well-being

Movie recommender systems are prevalent on streaming platforms, personalising content to enhance user engagement. However, there are concerns about their potential drawbacks. This paper investigates whether these systems can promote user well-being by integrating well-being metrics alongside traditional user behaviour data. We propose a Complementary Recommender System (CRS) framework that includes user well-being in the recommendation process. A quantitative experiment assessed the CRS's impact on various well-being components defined by the PERMA model. The CRS increased positive emotions, compared to behaviour-based systems. Engagement levels were similar across groups, suggesting the need for longer-term studies. The CRS led to greater improvements in relationship satisfaction and sense of achievement. Additionally, the CRS contributed to an increased sense of meaning, compared to the control group, partially validating its potential to positively influence well-being. These findings suggest that future algorithms can be designed to not only maintain user engagement but also contribute to their well-being in the digital age.

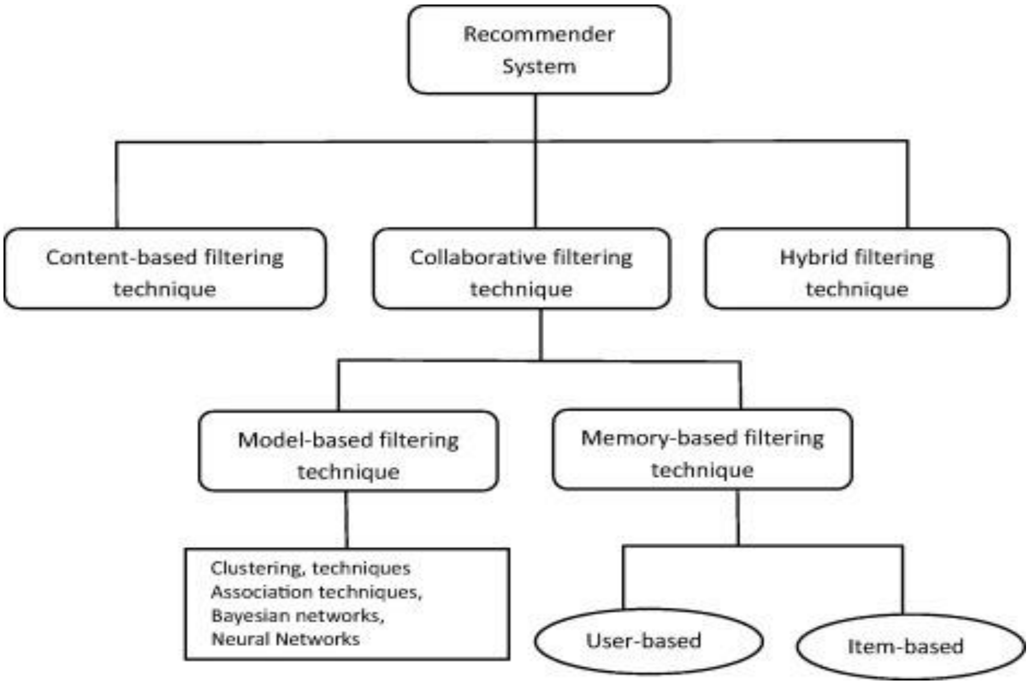
Keywords: recommender system, streaming platforms, well-being, PERMA

Biavaz, M., and Denegri-Knott, J. 2024. You are more than what you click: exploring the impact of movie recommender systems on consumer well-being. *Journal of Promotional Communications*, 11 (1), 22-53

INTRODUCTION

The landscape of media consumption has undergone a dramatic shift in recent years. The proliferation of streaming platforms has revolutionised the way we consume entertainment media, it has become "an environment where content is available anytime (instantaneousness), anywhere (ubiquity) and on all devices (mobility) has changed users' consumption habits and their relationship with content and content providers" (Gómez-Aguilar 2020, p.339). A recent report by Mintel (2023) indicates that streaming video has become mainstream for many internet users in the UK. These platforms leverage algorithms to suggest content tailored to individual users' preferences, and have significantly impacted how consumers make decisions, find information and content. These algorithms personalise content and product suggestions based on user behaviour and preferences. (Banker and Khetani 2019).

Figure 1: Different Recommendation Techniques (Isinkaye et al. 2015)



Companies are using different types of recommender systems that personalise product or content recommendations to users (see figure 1). Regardless of their type, recommender systems benefit both platforms and consumers: these algorithms streamline content discovery and provide personalised recommendations that enhance their viewing experience and platforms leverage them to increase user engagement, measured by factors like watch time and clicks, ultimately leading to higher user retention. They excel at keeping us glued to our screens by meticulously analysing our viewing habits and preferences, however, this relentless pursuit of engagement raises concerns about potential downsides. While recommendation systems promise to enhance user experience and content discovery, their impact on human lives is questionable, therefore several studies have been dedicated to the consequences of these algorithms. A key driver of risk comes from the way a service optimises its recommender systems for greater engagement. If it operates on an advertising-based business model, it has an incentive to increase user engagement – and in particular time online – to grow its revenue (eSafety Commissioner 2022).

Engagement-driven algorithms often prioritise content that keeps users glued to their screens, potentially creating negative consequences. One major concern lies in the creation of "filter bubbles". Recommender algorithms often prioritise content similar to what we've already watched or shown interest in, therefore, users are primarily exposed to content that aligns with their existing preferences. While this can be convenient and reduce the effort users need to search (Ansari, Essegaier and Kohli 2000), it restricts our exposure to diverse perspectives, viewpoints and contents (Nguyen et al. 2014) . Over time, it can lead to intellectual stagnation and limit our ability to engage with challenging or unfamiliar ideas. This promotes confirmation bias – "the tendency to seek and believe information aligning with one's existing opinions" (Henry et al. 2023). Present bias, a psychological concept describing the tendency to prioritise immediate gratification over potential long-term benefits when making a decision, presents another issue stemming from personalised content recommendation.

Present bias affects various decisions, such as snack choices, retirement savings, and movie selection. For instance, people often opt for enjoyable but forgettable movies like "The Fast and the Furious" for immediate gratification, whereas they are more likely to choose intellectually stimulating films with lasting value, like "Schindler's List," for future viewing (Lukoff 2021). Hence, we may add movies we consider important or enriching ("should watch" movies) to our queue, but prioritise watching more entertaining or enjoyable options ("want to watch" movies). Research by Harvard and the Analyst Institute confirms this behaviour, highlighting how users often neglect "should watch" movies in favour of more immediately appealing choices (Pariser 2011, p.117). Meantime, recommender systems (RS), can exacerbate this present bias as they primarily rely on behaviourism, selecting recommendations based on user actions ("what users do") while largely neglecting explicit preferences ("what users say") (Ekstrand and Willemsen 2016). As user actions reflect present bias, RS will prioritise items offering short-term gratification over long-term value. Consequently, RS may reinforce immediate desires rather than assisting users in achieving their long-term goals such as reaching their ideal self (Lyngs et al. 2018).

Movie recommender algorithms can, however, flip the script. While they can create the afore mentioned challenges, they also hold potential to contribute positively to consumer well-being, extending beyond simply keeping people engaged. This is because people don't always watch movies or TV shows solely for entertainment, they can serve more serious purposes related to mental wellness and overall well-being. Watching a movie can be a particularly positive experience, with the potential to impact the viewer's emotions (Busselle and Bilandzic 2009; Cupchik 2011; Green and Brock 2000), cognitive processes (Mares and Cantor 1992), and even behaviour (Chu et al. 2020). Further supporting this notion is the growing popularity of cinematherapy as a therapeutic tool for individuals with mental illnesses. It involves therapists prescribing movies to encourage beneficial modelling behaviours in patients (Mangot and Murthy 2017).

In the context of this research, well-being is defined using Martin Seligman's PERMA model (2011). This theory posits that well-being is influenced by these components: Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment. Also, through our primary research, we will answer the question: Can recommender algorithms in movie streaming platforms be optimised to contribute to consumers' well-being? While there is a wealth of research on the impact of recommendation systems and subjective well-being (Hemans and Wiredu Ocansey 2021), the number of studies that have looked directly at the effects of recommendation systems on subjective well-being is not adequate, and when investigating recommender algorithms and well-being, we can find only a few empirical research and validating data to understand how RS are negatively impacting users' well-being (Metzler et al 2022; Pattrn.ai 2023; CMA 2021). Having said that, all of this research, to the best of our knowledge, has focused on social media algorithms and none of these endeavours have explored specifically the potential effects of movie recommender algorithms on their users' wellness.

Therefore, addressing this research gap is crucial. Given the limited research on movie recommender systems and well-being, there's a significant opportunity as by optimising these systems, we can potentially harness the power of movies to serve as a promising intervention that contributes positively to user well-being. We propose a framework for a complementary recommender system (CRS) that integrates users' current well-being status with their viewing history to generate recommendations intended to promote and enhance their overall wellness, as opposed to traditional personalisation techniques that only rely on users' past behaviour. By evaluating the impact of this framework on user well-being, this research seeks to determine if such a system can contribute to consumers' well-being. This paper builds upon existing research on recommender systems and user well-being by proposing a novel Complementary Recommendation System (CRS). This system makes two key contributions to existing knowledge: The CRS framework incorporates well-being components alongside traditional user behaviour data. This approach moves beyond engagement-focused algorithms and provides empirical data on the impact of well-being-aware recommendations within movie streaming platforms. This data can be used in future research and design principles for recommender systems that prioritise user well-being alongside engagement. Secondly, it explores the concept of "therapeutic value" in movies and contributes to a richer understanding of how film consumption can be leveraged for boosting users' wellness.

LITERATURE REVIEW

Recommender Systems

Figure 2: Content-Based RS (Ng 2024)

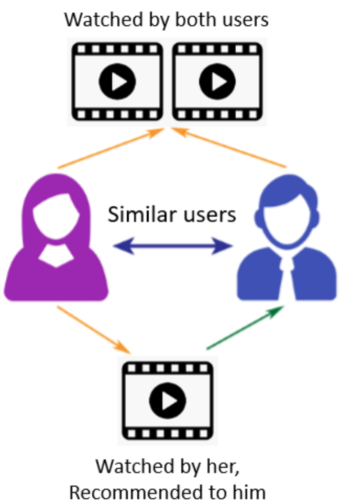
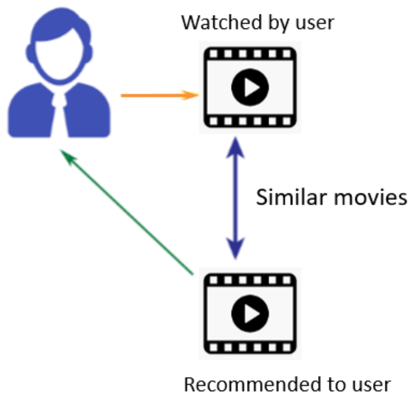


Figure 3: Collaborative Filtering (Ng, 2024)



Streaming platforms evaluate various factors to estimate the likelihood of a user engaging with specific titles within Netflix's catalogue. These factors include the user's interactions (such as viewing history and ratings), preferences of similar users, and characteristics of the items themselves (e.g. genre, actors) viewing time, and device usage (Netflix 2022). Recommender algorithms evolve over time, with user actions superseding initial preferences. Continuous improvement is driven by user feedback, which informs algorithm refinement to deliver fresh recommendations, aiming to enhance user satisfaction and enjoyment (Netflix 2022). There are various recommendation techniques which are being used on streaming platforms to generate personalised recommendations. Content-based filtering techniques provide recommendations to users based solely on individual user behaviour and data. These recommendations rely on item descriptions and specific user profiles (Thorat, Goundar and Barve 2015, as cited in Patel, Desai and Panchal 2017)

The underlying principle of content-based recommender systems involves two key steps. First, analysing the descriptions of items that can be used to characterise them, and storing these preferred characteristics in user profiles. Second, comparing the attributes of each item with the user profile to recommend items that exhibit a high degree of similarity to the user profile (J. Lu et al. 2015, as cited in Patel, Desai, and Panchal 2017). In collaborative filtering, recommendations for active users are generated by comparing their preferences and interests to those of similar users who have rated similar items (Thorat et al. 2015). Collaborative filtering systems review multiple common items to generate sets of users, which influences recommendation results. Applications of collaborative filtering typically involve very large datasets. Ultimately, hybrid filtering aims to provide better and more effective recommendations than a single algorithm by leveraging the strengths of each to address cold-start, sparsity, and scalability issues (Patel, Desai, and Panchal 2017).

Advantages of Recommender Systems

Advantages of recommender systems are well-documented. By and large the different methodological developments which we explained in this chapter, have aimed to yield more accurate recommendations believed to improve consumer outcomes by reducing search costs, increasing decision quality, and delivering greater satisfaction (Häubl and Trifts 2000; Xiao and Benbasat 2007). These technologies personalise interactions, fostering deeper customer engagement across diverse industries. Studies reveal that recommendations can account for a staggering 30% of revenue on major commercial platforms. Even a marginal improvement (1%) in recommendation quality can translate into billions of dollars. Furthermore, beyond the impressive revenue potential, companies leverage recommender systems for a multitude of reasons: enhanced customer retention, sales growth, and inventory management (NVIDIA 2024).

Disadvantages of Recommender Systems

This section explores negative impacts of the recommender systems on streaming platforms, mainly highlighting issues associated with users' choice making and content serendipity, that ultimately hinder the potential for movies to contribute to positive experiences, as seen in the concept of cinematherapy. Research conducted by Krook and Block (2023) explores the potential discrepancy between user engagement facilitated by recommender systems and the formation of authentic desires. It builds upon the concept of hierarchies of desires proposed by philosophers like Harry Frankfurt and Gerald Dworkin. As we explained earlier, recommender systems on platforms like Netflix, personalise user experiences by suggesting products or content based on algorithmic analysis of past behaviour and their ratings. While these systems claim to offer users what they "want," the study argues that they might primarily cater to immediate desires as reflected in past choices, potentially undermining the formation of second-order desires – conscious reflections on what individuals truly want to watch. The study further suggests that recommender systems might rely on a limited form of user engagement, characterised by passive consumption rather than active seeking that can negatively affect user autonomy and control.

This is like what Pariser (2011) explains in his book as present bias, a psychological tendency when people prioritise immediate gratification over potential long-term advantages when making decisions. Users may add movies they deem important or educational ("should watch" movies) to their watchlist but tend to watch more enjoyable or entertaining options ("want to watch" movies) instead. Research by Harvard and the Analyst Institute supports this, showing that users often ignore "should watch" movies in favour of more instantly gratifying choices (Pariser 2011, p.117). Moreover, over time algorithms tend to perpetuate limited diversity by relying on past user preferences and consumption patterns which results in the isolation of user perspectives (Nguyen et al. 2014). By prioritising content similar to that which users have previously engaged with, algorithms may overlook diverse genres or niche interests, contributing to a homogenised content landscape which results in filter bubbles. The filter bubble refers to the idea that “recommenders isolate us from a diversity of viewpoints, content, and experiences, and thus make us less likely to discover and learn new things” (Pariser 2011 cited in Chu et al. 2020).

Another flaw in algorithms is their inherent bias, which often manifests in the promotion of platform-owned content over independent or lesser-known works. The article (missing citation) illustrates this phenomenon by highlighting the case of "The Sun," a Taiwanese crime epic widely acclaimed internationally but overlooked by algorithms due to lack of promotion. Consequently, users may miss out on high-quality content that does not align with mainstream trends or the platform's profit motives (Martinez 2021). Also, the oversimplification of content into broad categories on streaming platforms has made it challenging for niche and content-rich shows to get through platform filters, leading to the rise of what is termed "Ambient TV" (lacking citation) which refers to content that is designed to produce minimal provocation of thoughts or feelings in viewers, serving primarily as background noise. This potentially leads to the proliferation of content that lacks depth and complexity (Martinez 2021). As discussed, previous research on recommender systems has primarily focused on negative outcomes, particularly the limitation of content discovery and user agency in choice-making. However, the potential impact on well-being has received less attention, despite the acknowledged influence of movies on wellness and the concept of cinematherapy that happens through a combination of enjoyment, focused attention, and a sense of empowerment. We experience pleasure while watching, become fully engaged in the story, and are sometimes left feeling inspired or gaining new insights into ourselves, others, or the human condition as a whole (Berg-Cross, Jennings, and Baruch 1990 cited in Niemiec 2020). To address the afore mentioned gap, a clear definition of well-being within this context is necessary.

Defining Well-being

There are many ways to define and measure well-being. Well-being can be looked at from an objective standpoint (like having enough resources to survive, access to education, or clean air) or a subjective one (how we feel about our lives) (Forgeard et al. 2011). Psychologists have been studying this for a while. For instance, Diener (1984) pointed out that subjective wellbeing involves affective and cognitive elements. Ryff and Keyes (1995) came up with six areas that contribute to psychological well-being: self-acceptance, positive relationships with others, autonomy, environmental mastery, purpose in life, and personal growth. Keyes (2002) suggests that to truly flourish, you need to feel good emotionally, psychologically, and socially. There is another well-being model that has been introduced by Saligman (2011). Saligman’s well-being theory is known as the PERMA model (which this research utilises to measure well-being). In Seligman’s theory, well-being is defined as a combination of cognitive happiness, hedonic happiness, and eudaimonia (Forgeard et al. 2011). According to this model, well-being is predicted by five elements: Positive Emotion (P), Engagement (E), Relationships (R), Meaning (M), and Accomplishment (A). In this theory, each element that contributes to well-being can be pursued for its own sake and is independently defined and measured (Seligman 2011).

Well-being may be increased by increasing PERMA elements. Positive Emotion includes subjective reports of happiness, hope, joy, and satisfaction. Engagement is an element that represents flow, and it refers to focus, interest, or absorption in an activity. Relationships include closeness and connection with family, friends, or colleagues. These relationships are important throughout a person's lifespan and contribute to well-being in many ways. Meaning is belief in, or membership of, something larger than oneself, and may be derived from religion, spirituality, or advocacy. The final element, Accomplishment, refers to pursuits that occur throughout life for the sake of 'winning.' Accomplishment often requires perseverance and resilience, and may include academics, athletics, or career achievements (Seligman 2011).

While numerous models explore well-being, Seligman's PERMA model offers a particularly well-suited framework for analysing the impact of movie streaming platforms on user well-being. This model was chosen because it goes beyond just happiness and considers five key elements that contribute to well-being, and movies can portray and even generate these PERMA elements in viewers. (Niemiec and Wedding 2014, p.352). Also, PERMA is a core model in positive psychology, which studies happiness and flourishing. Since movies can evoke positive emotions, promote engagement, and offer meaning, PERMA is well-suited for analysing these effects (Niemiec and Wedding 2014, p.352).

Movies and Well-being

Movies are subjective and complex products and could have therapeutic value if they are recommended at the right time. However, regarding streaming services, recommended algorithms primarily address the "engagement" element of the PERMA model, and they overlook or potentially undermine other elements as they are unaware of which element a specific user is lacking in their life. Having said that, alongside fostering engagement (E), movies are uniquely positioned to not only depict positive emotions (P) and meaning (M) in the viewer but also to generate them (Niemiec and Wedding 2014, p.352). Of course, this will depend on the type of film shown, the viewer's level of identification with the characters, and many other factors. People probably do not go to the movies to observe these phenomena on the screen, rather they are far more likely to head to the theatre, because the film might lead them to feel positive emotions, engage fully with the material, and experience meaning. Numerous films illustrate positive emotions, engagement, and meaning (Niemiec and Wedding 2014, p.352). Films can serve as subtle tools for teaching exemplary behaviours and have been incorporated into therapeutic practices, a field known as cinematherapy (Hesley and Hesley 2001 cited in Popa et al. 2021). The therapeutic use of films dates to the early 20th century (Powell & Newgent 2010), particularly in military psychiatric hospitals during the 1940s. Katz (1945) noted that films were used not just for entertainment, but also for educational, vocational, and inspirational purposes in a Social Therapy Program for Neuropsychiatry. Whitmyre (1958) observed that incorporating motion pictures had become an accepted part of treatment programs for hospitalised psychiatric patients.

In a survey done by Lampropoulos, Kazantzis, and Deane (2004) 827 licensed practising psychologists were included to determine whether professional psychologists use motion pictures in their clinical practice. The results revealed that 67% reported the use of motion pictures to promote therapy gains and 88% considered the use of motion pictures as effective in promoting treatment outcome. Cinematherapy has been successfully utilised with diverse client groups, including adolescents, couples, and older adults, to facilitate emotional expression, gain insights, and develop coping skills (Dermer and Hutchings 2000; Bierman et al. 2003; Marsick 2010). Furthermore, outside a clinical context, movies can boost our well-being by engaging viewers emotionally, providing a safe space to explore and release difficult emotions, which is known as catharsis. The concept of catharsis, as noted by Aristotle in the context of Greek tragedies, is relevant to modern film therapy. Movies allow viewers to experience and purge emotions by empathising with characters, providing a therapeutic release like that experienced by audiences of ancient tragedies (Hamilton 2023). Also, research, utilising both qualitative and quantitative methodology (Grobler, 2012), demonstrated that films can significantly influence people's well-being. Participants watched positive psychology movies and were able to identify strengths depicted by the characters and reported using these strengths in their own lives. Quantitative results showed that viewing these films led to increases in certain strengths (e.g., humour, open-mindedness, kindness), general well-being, and emotional well-being. In conclusion, by incorporating diverse cinematic experiences into users' lives, individuals can enhance their overall happiness and fulfilment. Therefore, streaming platforms need to reform their systems and come up with a recommendation system that complements behaviour-based algorithms. This well-being-aware recommender system has been called “complementary recommender system (CRS)”.

THEORETICAL FRAMEWORK

To address the aim of this research, the objectives and hypotheses are proposed aligned to PERMA elements:

Positive emotions: Behaviour-based movie recommender systems focus on past user behaviour, potentially neglecting the promotion of positive emotions. This can lead to repetitive content recommendations, overlooking opportunities to introduce uplifting movies that enhance a user's current emotional state. For instance, a user consistently watching thrillers might only receive similar suggestions, missing the potential benefits of a comedy when feeling down. A meta-analysis showed that film clips can effectively induce various mood states (Fernández-Aguilar et al.,2019). Thus the proposed system can leverage this to recommend movies that evoke positive emotions and boost well-being. We hypothesise that the complementary recommender system will increase positive emotions more than behaviour-based systems.

Engagement: As explained in detail at the beginning of this chapter, behaviour-based recommender systems analyse and predict user preferences by examining past behaviour, and using algorithms to identify patterns and suggest movies that align closely with individual tastes. This approach can reinforce existing preferences, continuously recommending films that enhance user satisfaction and engagement. Our objective is to compare the effect of complementary and behaviour-based systems on engagement, hypothesising that behaviour-based algorithms better boost engagement.

Relationships: Observational learning suggests individuals can learn behaviours and attitudes by observing others, including those in movies. Movies offer viewers models of behaviour and comfort by showing relatable experiences, potentially validating or challenging their perceptions (Heston and Kottman 1997; Solomon 1995, 2001 cited in Lampropoulos, Kazantzis and Deane 2004). By observing how characters navigate relationships, viewers may gain insights into their own behaviours and attitudes within their relationships. Heston and Kottman (1997) presented a case study showing how a client with strained familial relations, improved her depressive symptoms and communication with her mother after watching a film depicting intergenerational relationships (Lampropoulos et al. 2004).

Movies offer considerable potential for therapeutic interventions across various contexts, including individual, couples, family, and group therapy. They can effectively address a wide array of presenting issues such as familial conflicts, trauma, relationship dynamics, to name a few (Hesley and Hesley 1998; Solomon 1995, 2001 cited in Lampropoulos et al. 2004).

A complementary system can recommend movies with meaningful portrayals of relationships, promoting positive changes in viewers' attitudes and behaviours. We hypothesise that it will improve users' perceptions of relational quality more than behaviour-based systems.

Meaning: Algorithms may undermine users' sense of meaning due to the echo chamber and filter bubbles. In other words, the vast amount of available data may lead to users experiencing intellectual isolation. Website algorithms, drawing from users' past interactions, selectively predict the information they are likely to find engaging. As a result, users may find themselves isolated from information that diverges from their existing viewpoints, perpetuating their immersion in cultural and ideological bubbles (Haddad 2022). In addition, Noordeh et al. (2020) reveal that extended exposure to recommendations generated by systems significantly reduces the diversity of content, guiding individual users into echo chambers defined by a limited range of material. Hence, due to this limited content discovery and serendipity, users may miss out on encountering content that challenges their beliefs or expands their understanding of the world. Therefore, we hypothesise that current recommender algorithms negatively impact users' perception of meaning, while a complementary system can positively contribute. We will measure and compare their impacts on users' sense of meaning.

Accomplishment: As we discussed, traditional streaming platforms often prioritise users' immediate desires, known as first-order desires, which typically revolve around seeking entertainment or instant gratification. However, this approach may overlook users' deeper aspirations and values, termed second-order desires, which encompass broader objectives such as personal growth or emotional well-being. This failure to consider second-order desires can potentially lead to a loss of autonomy for users.

According to philosopher Harry Frankfurt, human autonomy derives from our capacity to form and prioritise second-order desires, allowing us to exert control that transcends our initial impulses. Gerald Dworkin extends this concept by positing that autonomy is achieved when individuals reflect on their desires, evaluate them rationally, and adopt new preferences as their own. A complementary system can enhance autonomy by recognising and prioritising users' second-order desires, offering content that supports broader objectives like personal development. This encourages users to engage in reflective thinking, expanding their choice architecture and providing access to diverse content options.

Research conducted by Guo (2018), investigates how algorithms could generate recommendations for self-actualisation, which is important to achieving psychological well-being by continuing to grow and develop throughout life (Main 2023). This research proposes a novel approach, current systems can restrict users to a narrow range of information by filtering out items deemed unlikely to be liked. This creates a "filter bubble" that limits exposure to diverse viewpoints. The proposed approach tackles this issue by offering users additional recommendation lists alongside the traditional Top-N suggestions. These lists aim to expand users' horizons and potentially challenge their preferences. For instance, one list might include items the system predicts the user will dislike, allowing them to discover unexpected favourites. By providing these diverse options, recommender systems can empower users to broaden their horizons and discover new things that resonate with them. In line with Guo's contribution, our proposed system, CRS, aims to lead to more rewarding content consumption. We hypothesise that it will increase the sense of achievement more than behaviour-based systems.

METHODOLOGY

This research used a quantitative methodology to compare the effects of complementary and behaviour-based recommender systems on users' well-being. We employed a quasi-experimental design, incorporating mainly quantitative elements. "Quasi-experimental study designs, often described as non-randomised, pre-post intervention studies" (Harris et al.

2006.16), are well-suited for research settings where random assignment of participants to experimental conditions is not feasible or ethical, and a quasi-experimental design seeks to ascertain a causal connection between independent variable and dependent variable (Reichardt 2019). Given the nature of this study, which involves comparing the effects of different movie recommendation systems (independent variable) on users' well-being (dependent variable), a quasi-experimental design allows for controlled comparisons between groups while accounting for pre-existing differences among participants.

In our case, randomly assigning participants to different movie recommendation groups would not be feasible, as it would neglect the personalised nature of the recommendations. By employing a quasi-experimental design, we can assign participants to groups based on specific criteria related to their well-being state, ensuring that the intervention is relevant and meaningful to each individual. The study involved 45 participants (female: 24, male: 18, non-binary: 3) recruited from diverse backgrounds. Participants were at least 18 years old and interested in watching movies. Researchers used a multi-faceted approach for recruitment, including social media platforms and encouraging referrals from existing participants. Participants were divided into three groups of 15 each: complementary recommender system, behaviour-based recommender system, control group. Additionally, quasi-experimental design was employed to allow for a realistic yet controlled assessment of how different recommendation strategies influence users' well-being. Although participants in the complementary recommender system group were required to watch a specific movie based on their well-being profiles, this mimics real-world scenarios where users are encouraged to engage with certain curated content. Similarly, the traditional recommender group interacted with platform-generated recommendations, reflecting typical viewing experiences. By preserving the natural viewing environments while controlling movie selection across groups, the design ensures ecological validity, capturing how recommendation systems operate in practice and their impact on users' well-being.

The primary instrument used for data collection is a structured questionnaire administered to participants before and after movie viewing. The questionnaire includes items adapted from the PERMA profiler, a well-established tool for measuring psychological well-being across five dimensions: Positive Emotions, Engagement, Relationships, Meaning, and Achievement. In the pre-questionnaire, validated scales were used to assess participants' baseline well-being before any intervention. These questions have been called Core Well-Being Measures and are formulated to evaluate participants' current state of well-being, enabling the personalisation of movie recommendations accordingly. Additionally, participants were asked about the impact of movies they had recently watched on their well-being. These questions have been called Tailored Intervention Impact Measures and established a baseline for their usual movie-watching experience. In the post-questionnaire, the same well-being scales (Core Well-Being Measures) were used to assess participants' well-being at phase 2. Also, participants were asked about the impact of the specifically recommended movie (intervention) on their well-being (Tailored Intervention Impact Measures). These questions allowed us to measure the targeted effect of the recommendation system. Before the experiment, participants were informed about the study objectives and gave consent. They completed a pre-experiment questionnaire assessing their baseline well-being and recent movie impacts. The data showed a normal distribution for all well-being elements except engagement. (see figures 2 to 5).

Figure 4:
Histogram of P Score

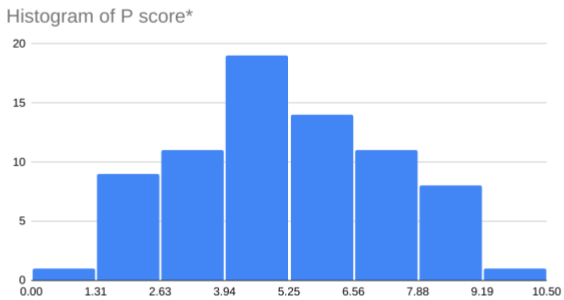


Figure 5:
Histogram of R Score

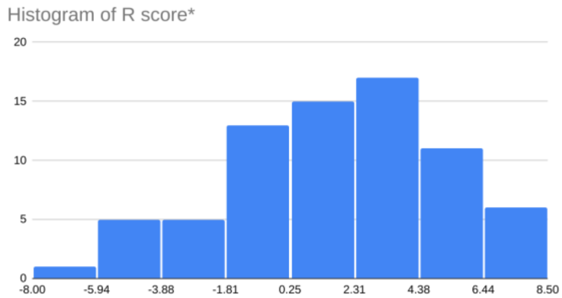


Figure 6:
Histogram of M Score

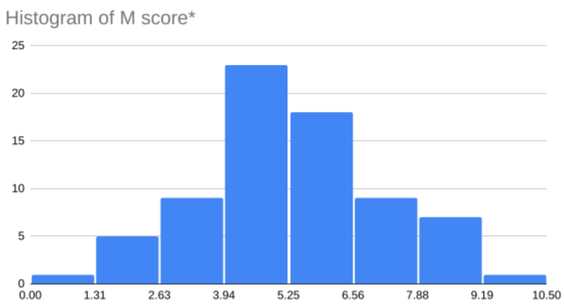
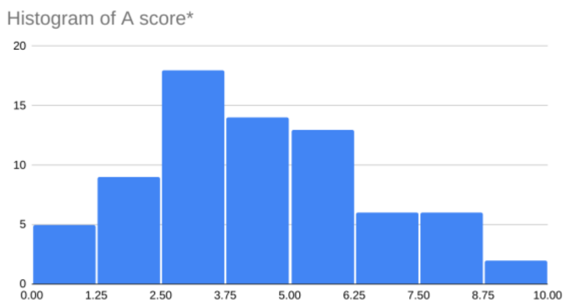


Figure 7:
Histogram of A Score



This was also confirmed using the Kolmogorov-Smirnov (KS) test in Python.

Utilising the concept of central tendency, we calculated the mean for each well-being element. By applying the empirical rule (approximately 68% of the data falls within one standard deviation from the mean), we selected a sample of individuals who comprise the first 16% of our population, which is actually the lower bound of the spread of the data (see Figure 8). It represents the value below which 16% of the data points fall. For instance, in the context of positive emotions (P score), where the mean score was calculated as 5.2 with a standard deviation of 2.1, individuals scoring below 3.1 were identified as needing a movie recommendation aimed at enhancing their positive emotions. Likewise, we identified participants who need a specific well-being component. For engagement, due to its non-normal distribution to determine that 16% for "engagement," we simply find the 16th percentile. In a non-normally distributed dataset, the 16th percentile provides insight into the dispersion of the data's lower tail, and in our study, it is approximately 3 (Figure 9). By leveraging these statistical techniques, we were able to assign participants to one of the three experimental groups, and tailor movie recommendations to individuals based on their unique well-being profiles.

Participants in the complementary recommender system group were exposed to movie recommendations generated by the complementary recommender system, and tailored to their individual well-being profiles. To be more precise, movie selection for this group was based on Core Well-Being Measures, specifically targeting the PERMA element that each participant needed the most. The films were selected from "*Positive Psychology at the Movies: Using Films to Build Character Strengths and Well-Being*" by Niemiec and Wedding (2013), which examines the intersection of positive psychology and cinema. This book discusses nearly 1,500 films that exemplify positive psychology themes, illustrating how characters and narratives can represent virtues and strengths, and provide inspiration for personal development. Additionally, the authors explain the PERMA well-being model and explore how its elements are reflected across various films.

Participants in the behaviour-based recommender system group were told to watch one of the recommended movies on the streaming platform they were actively using, therefore they received recommendations based on their past viewing behaviour and preferences. The control group received random movie recommendations without consideration of their well-being profiles. After watching the recommended movies, participants completed the post-experiment questionnaire, which mirrored the pre-experiment questionnaire. To assess the impact of different recommendation systems on users' well-being and movie preferences, inferential statistical analyses were performed. In our data analysis, we utilised the mixed-design analysis of variance (ANOVA) to investigate potential differences among the three groups in each component of the PERMA model. A mixed model ANOVA blends elements of both between-subjects and within-subjects ANOVA. It necessitates a minimum of two categorical independent variables, often referred to as factors, with at least one variable varying between subjects and another variable varying within subjects. (Murrar and Brauer 2018) In our case, we had two independent variables:

Time: Time serves as a within-subjects factor because each participant experiences both time points.

Movie Recommendation System: Participants are assigned to only one of these groups, hence it's a between-subjects factor.

By and large, we conducted within-subjects analysis to compare participants' well-being scores before and after the intervention (watching the movie) within each group. Additionally, we conducted between-subjects analysis to compare well-being scores among the three groups to determine if there are any differences in well-being outcomes based on the type of movie recommendation system they received. Furthermore, post hoc test (Tukey's Honestly Significant Difference) was essential for our analysis to compare specific pairs of groups to determine where the differences lie.

Validity

According to Reichardt and Little (2019), one of the threats to validity is history effects and in the context of a pre-test-post-test design, history effects refer to the influence of external events that occur between the pre-test and post-test measurements. These external events, separate from the treatment being studied, can impact the outcome observed at the post-test stage. In our exploration, some participants experienced a gap of approximately one week between the pre-test and post-test assessments, while others faced up to a two-week gap. Therefore, we acknowledge the susceptibility of our study to external factors that may affect participants' mood, behaviour, and receptiveness to movie recommendations. Factors such as work-related stressors, personal events, or changes in social dynamics could inadvertently impact participants' well-being and their engagement with different types of movie recommendation systems. Throughout our study, we have implemented a measure to consider any influence of historical events on our outcomes, which was comparing the same mood-related questions in our pre and post questionnaires through running a paired t-test to detect any significant changes in participants' mood. This statistical method is suitable for evaluating whether there are significant differences in the means of two phases for the same participants.

As seen in tables 1 to 3, the lack of significant differences in Positive Emotions and Meaning, indicate that these dimensions were not notably influenced by external events. Therefore, any observed changes in these areas can be more confidently attributed to the

intervention of the movie recommender system rather than to extraneous factors.
Conversely, the significant difference observed in the Relationships dimension suggests that participants experienced changes in their sense of relationships during the study period. Consequently, this result indicates a potential history effect, which poses a threat to the internal validity of the study.

Table 1: Paired T-Test on M (Meaning) Score to Compare pre and Post Experiment

Paired Samples Test

		Paired ... 95% Confidence Interval of the ...			Significance	
		Upper	t	df	One-Sided p	Two-Sided p
Pair 1	pre_m - post_m	.336	-.501	44	.309	.619

Table 2: Paired T-Test on P (Positive Emotions) Score to Compare Pre and Post Experiment

Paired Samples Test

		Paired ... 95% Confidence Interval of the ...			Significance	
		Upper	t	df	One-Sided p	Two-Sided p
Pair 1	pre_p - post_p	.420	-.101	44	.460	.920

Table 3: Paired T-Test on R (Relationship) Score to Compare Pre and Post Experiment

Paired Samples Test

		Paired ... 95% Confidence Interval of the ...			Significance	
		Upper	t	df	One-Sided p	Two-Sided p
Pair 1	pre_r - post_r	-.032	-2.136	44	.019	.038

By adopting these strategies, we aimed to enhance the robustness of our findings and provide meaningful insights into the relationship between movie recommendations and consumers' well-being.

Another potential threat to internal validity was maturation, referring to natural changes in individuals over time (Mark and Reichardt 2001). It is crucial to differentiate this from history effects, which stem from external events impacting participants. Maturation effects arise internally due to time passage (Reichardt and Little 2019, p.103). To address this we compared identical mood-related questions to detect changes in participants' mood between pre-test and post-test, independent of the movie recommendation intervention. The 1-3 week study timeframe helped mitigate this threat.

Testing is another threat to internal validity where merely observing or measuring individuals can alter their behaviour (Reichardt 2019). This phenomenon can make people self-conscious, suspicious, or change their work ethic. Over time, individuals may get used to being observed or learn from the process. For example, in physical fitness studies, participants might perform better on post-assessments due to knowledge gained from initial assessments. In attitude studies, surveys can prompt reflection, altering attitudes in subsequent assessments. These behavioural changes can bias study outcomes, independent of actual treatment effects. To mitigate this "Testing" effect, we implemented blinding. Blinding minimises threats to internal validity (Page and Persch 2013) and maximises result validity (Karanicolas, Farrokhyar, and Bhandari 2010). Therefore, we kept participants unaware of the specific hypotheses, or the group to which they belong, to reduce the likelihood of altered behaviour due to observation.

Instrumentation was our fourth threat to internal validity which occurs in a pretest-posttest design when different measurement tools are used, biassing the treatment effect (Reichardt and Little 2019, p.104). In our research, we addressed this by ensuring consistent measurement instruments for both pretest and posttest assessments. Using identical questions and methods maintained the integrity and validity of our comparisons

In a pretest-posttest design, participant changes over time can bias the treatment effect estimation. Attrition, or experimental mortality, occurs when participants leave the study, while augmentation involves adding new ones (Reichardt and Little 2019, p.105). If pretest participants differ from posttest ones, it threatens internal validity. For instance, in a depression study, if the most depressed drop out, effectiveness might be overestimated (Reichardt and Little 2019, p.104). In our study, 27 out of 76 participants dropped out, leaving only 49, with data from 45 used, posing a threat to internal validity. Future studies should aim to retain participants throughout the entire study period by employing strategies such as incentives.

Reliability

In this research, we employed a combination of approaches due to the nature of our pre-post design, which included core well-being measures and tailored intervention impact questions, to ensure consistency in our findings.

Core Well-Being Measures: For the core well-being questions in the pre- and post-questionnaires for all three groups (complementary recommender, behaviour-based recommender, and control), Cronbach's Alpha was calculated to assess internal consistency. This statistic measures how closely related items are within a scale (UCLA 2021). A higher coefficient indicates better internal consistency. The results show moderate internal consistency for Groups 1 and 2, which is acceptable. Group 3 has a lower, but still acceptable coefficient (Hajjar 2018), possibly due to the greater heterogeneity of the control group (see Tables 4, 5, and 6).

Table 4: Cronbach's Alpha for Groups 1

Case Processing Summary

		N	%
Cases	Valid	15	100.0
	Excluded[a]	0	.0
	Total	15	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.725	5

Table 5: Cronbach's Alpha for Groups 2

Case Processing Summary

		N	%
Cases	Valid	15	100.0
	Excluded[a]	0	.0
	Total	15	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.665	5

Table 6: Cronbach's Alpha for Groups 3

Case Processing Summary

		N	%
Cases	Valid	15	100.0
	Excluded[a]	0	.0
	Total	15	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.575	5

Tailored Intervention Impact Measures: To assess the reliability of tailored items in our questionnaire, we used the Intraclass Correlation Coefficient (ICC) for Group 1 (complementary RS). ICC measures the consistency of ratings by comparing the variability of different ratings of the same subject to the total variation across all ratings and subjects (LIU et al. 2016). For Group 1, where participants rated a set of pre-selected movies (n = 5), the ICC analysis assessed the internal consistency of these ratings on well-being. The behaviour-based group could not use ICC due to the heterogeneity in movie selection, which undermines ICC's assumption of consistency in stimuli. Also, it is worth mentioning that the sample size (3 participants) is a limitation, and a larger sample would provide more robust estimates of consistency within group 1. For "Ferris Bueller's Day Off" (1986), the ICC of .790 indicates moderate consistency among the participants' ratings. For "Limitless" (2011), the ICC of .362 shows low consistency, likely due to the varied personal preferences for its themes. Limitless is a science fiction thriller that explores themes of ambition, power, and the potential dangers of cognitive enhancement. Some might see the movie as inspiring and empowering, while another might not, some people might be very interested in this genre, some might prefer more realistic movies. Away We Go" had an ICC of .638, indicating moderate agreement on its well-being impact. "About Schmidt" had a very low ICC of -1.950, showing poor reliability due to diverse personal reactions. "About Schmidt" likely evoked varied personal reactions, influencing participants' ratings and causing high variability. One out of three participants had a very positive experience, while the other two did not. This difference, combined with the small sample size, likely led to the poor reliability result, as even one diverse response can impact consistency. "Julie & Julia" had a

Table 7: ICC for movie Ferris Bueller's Day Off

Intraclass Correlation Coefficient

	Intraclass Correlation	95% Confidence Interval		F Test with True Value	
		Lower Bound	Upper Bound	Value	df1
Single Measures	.557	.022	.932	4.765	4
Average Measures	.790	.062	.976	4.765	4

Table 8: ICC for Movie Limitless

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value	
		Lower Bound	Upper Bound	Value	df1
Single Measures	.159[a]	-.299	.813	1.566	4
Average Measures	.362	-2.226	.929	1.566	4

Table 9: ICC for Movie Away We Go

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value	
		Lower Bound	Upper Bound	Value	df1
Single Measures	-.283[a]	-.451	.405	.339	4
Average Measures	-1.950[c]	-13.905	.671	.339	4

Table 10: ICC for Movie About Schmidt

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value	
		Lower Bound	Upper Bound	Value	df1
Single Measures	.491[a]	-.083	.919	3.897	4
Average Measures	.743[c]	-.297	.971	3.897	4

Table 11. ICC for Movie About Schmidt

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value	
		Lower Bound	Upper Bound	Value	df1
Single Measures	.370[a]	-.178	.888	2.765	4
Average Measures	.638[c]	-.828	.960	2.765	4

good ICC of .734, suggesting higher agreement on its well-being impact.

To ensure participants understood the research and made an informed decision, a written consent form was provided before participation, explaining the study's objectives, procedures, and participants' rights. Participants could withdraw at any point without penalty, with contact information available for assistance. Anonymity and data confidentiality were maintained

Table 13.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
time	Sphericity Assumed	17.778	1	17.778	8.466	.006
	Greenhouse-Geisser	17.778	1.000	17.778	8.466	.006
	Huynh-Feldt	17.778	1.000	17.778	8.466	.006
	Lower-bound	17.778	1.000	17.778	8.466	.006
time * group	Sphericity Assumed	14.022	2	7.011	3.339	.045
	Greenhouse-Geisser	14.022	2.000	7.011	3.339	.045
	Huynh-Feldt	14.022	2.000	7.011	3.339	.045
	Lower-bound	14.022	2.000	7.011	3.339	.045
Error(time)	Sphericity Assumed	88.200	42	2.100		
	Greenhouse-Geisser	88.200	42.000	2.100		
	Huynh-Feldt	88.200	42.000	2.100		
	Lower-bound	88.200	42.000	2.100		

throughout the research, with all data anonymised by removing personally identifiable information. A reference system linked pre and post-test results without including names, introducing a limitation discussed later. Several limitations may have influenced the study's outcomes. While the research provides valuable insights into the impact of recommender systems on user well-being, some limitations include: The sample size of 15 participants per group is considerably small for quantitative research. A larger sample would yield more robust and generalisable findings. An anonymised reference system was used to link pre and post-test results, protecting anonymity but causing issues. Some participants ignored the reference section in the pre-questionnaire or forgot their reference during the post-test, resulting in data that could not be analysed due to the lack of identifiable links between responses.

FINDINGS

Our study explored the effects of different movie recommendation systems on users' well-being. Key findings related to each wellbeing component are detailed below:

Table 12.

Descriptive Statistics

	group	Mean	Std. Deviation	N
pre_positive_emotions	complementary RS	4.73	2.154	15
	behaviour-based RS	5.67	1.877	15
	control	4.47	1.457	15
	Total	4.96	1.882	45
post_positive_emotions	complementary RS	6.60	1.404	15
	behaviour-based RS	5.60	1.765	15
	control	5.33	1.633	15
	Total	5.84	1.665	45

Positive Emotion State

Descriptive statistics for pre- and post-questionnaire scores on positive emotions are shown in Table 12. The complementary recommender system group had the highest average post-positive emotion score (6.60), followed by the behaviour-based recommender group (5.60) and the control group (5.33). Groups 1 and 3 saw increases of 1.87 and 0.84, respectively, while Group 2 decreased by 0.07.

A repeated-measures MANOVA revealed a significant main effect of time (F=8.466, Sig = .006) and a significant interaction effect between time and group (F=3.339, Sig = .045). This suggests different impacts of the recommender systems on positive emotions across the groups.

Table 14.

ANOVA

positive_emotions_change

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28.044	2	14.022	3.339	.045
Within Groups	176.400	42	4.200		
Total	204.444	44			

Table 15.

Post Hoc Tests

Multiple Comparisons

Dependent Variable: positive_emotions_change
Tukey HSD

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% ... Lower Bound
complementary RS	behaviour-based RS	1.93333*	.74833	.035	.1153
	control	1.00000	.74833	.383	-.8181
behaviour-based RS	complementary RS	-1.93333*	.74833	.035	-3.7514
	control	-.93333	.74833	.433	-2.7514
control	complementary RS	-1.00000	.74833	.383	-2.8181
	behaviour-based RS	.93333	.74833	.433	-.8847

Here is the result for between-subject effects of "group" by running a one-way ANOVA on positive emotions change (post-positive emotions score minus pre-positive emotions score) which shows a p value of .045, therefore we conducted a post hoc test (Tukey's HSD) afterwards to compare the mean differences in positive emotion change scores between specific groups.

The effect size ($\eta^2 = 0.137$) suggests a large effect based on conventional thresholds. Such a large effect size highlights the practical significance of the complementary recommender system in enhancing positive emotions, demonstrating that well-being-focused recommendations can have a substantial impact on users' emotional states. Tukey's HSD revealed significant differences in positive emotion change scores between the complementary RS and behaviour-based RS group (p=.035). Participants in the complementary RS group experienced a greater increase in positive emotions (mean difference=1.93) compared to the behaviour-based RS group.

Engagement Levels

As shown in Table 16, pre-questionnaire scores for engagement were similar across the

Table 16.

Descriptive Statistics

	group	Mean	Std. Deviation	N
pre_engagement	complementary RS	6.33	1.839	15
	behaviour-based RS	6.33	1.718	15
	control	5.47	2.200	15
	Total	6.04	1.930	45
post_engagement	complementary RS	5.40	1.882	15
	behaviour-based RS	6.87	1.457	15
	control	5.93	1.831	15
	Total	6.07	1.802	45

complementary and behaviour-based groups (6.33), with the control group scoring 5.47. Post-questionnaire scores showed a decrease in engagement for the complementary RS group (-.93) and an increase for the behaviour-based RS group. A within-subjects ANOVA did not reveal a significant main effect of time (F=.003, p=.953), suggesting unchanged engagement levels overall. The interaction effect between time and group was not statistically significant (F=1.598, p=.214). A between-subjects ANOVA also showed no significant differences in engagement changes scores between the groups. Moreover, the effect size (η^2 = 0.070) indicates a small effect size, suggesting that only a small proportion of the variance in engagement scores can be attributed to the independent variable.

Table 17.

ANOVA

engagement_change

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	20.578	2	10.289	1.598	.214
Within Groups	270.400	42	6.438		
Total	290.978	44			

Table 18

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
time	Sphericity Assumed	.011	1	.011	.003	.953
	Greenhouse-Geisser	.011	1.000	.011	.003	.953
	Huynh-Feldt	.011	1.000	.011	.003	.953
	Lower-bound	.011	1.000	.011	.003	.953
time * group	Sphericity Assumed	10.289	2	5.144	1.598	.214
	Greenhouse-Geisser	10.289	2.000	5.144	1.598	.214
	Huynh-Feldt	10.289	2.000	5.144	1.598	.214
	Lower-bound	10.289	2.000	5.144	1.598	.214
Error(time)	Sphericity Assumed	135.200	42	3.219		
	Greenhouse-Geisser	135.200	42.000	3.219		
	Huynh-Feldt	135.200	42.000	3.219		
	Lower-bound	135.200	42.000	3.219		

Table 19.

Descriptive Statistics

	group	Mean	Std. Deviation	N
pre_relationship	complementary RS	3.07	1.870	15
	behaviour-based RS	4.87	2.100	15
	control	4.93	2.282	15
	Total	4.29	2.222	45
post_relationship	complementary RS	5.47	1.552	15
	behaviour-based RS	4.60	1.502	15
	control	4.60	1.724	15
	Total	4.89	1.613	45

Impact on Relationships

As shown in Table 19, pre-questionnaire scores on relationship well-being varied across groups: the complementary RS group increased by +2.4, while the behaviour-based RS and control groups decreased by -0.27 and -0.33, respectively. The within-subjects ANOVA showed no significant main effect of time ($p = .170$), indicating no overall change in relationship well-being across all groups after using the recommender systems. However, a significant interaction effect between "time" and "group" was found ($p = .019$), suggesting different impacts of the recommender systems on relationship well-being across the groups. This indicates that while the average score across all groups remained relatively stable, the pattern of change differed among them.

Table 20.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F
time	Sphericity Assumed	8.100	1	8.100	1.945
	Greenhouse-Geisser	8.100	1.000	8.100	1.945
	Huynh-Feldt	8.100	1.000	8.100	1.945
	Lower-bound	8.100	1.000	8.100	1.945
time * group	Sphericity Assumed	36.467	2	18.233	4.378
	Greenhouse-Geisser	36.467	2.000	18.233	4.378
	Huynh-Feldt	36.467	2.000	18.233	4.378
	Lower-bound	36.467	2.000	18.233	4.378
Error(time)	Sphericity Assumed	174.933	42	4.165	
	Greenhouse-Geisser	174.933	42.000	4.165	
	Huynh-Feldt	174.933	42.000	4.165	
	Lower-bound	174.933	42.000	4.165	

Table 21.

ANOVA

meaning_change

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	52.578	2	26.289	3.383	.043
Within Groups	326.400	42	7.771		
Total	378.978	44			

Table 22.

Post Hoc Tests

Multiple Comparisons

Dependent Variable: relationship_change
Tukey HSD

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% ... Lower Bound
complementary RS	behaviour-based RS	2.66667*	1.05389	.040	.1062
	control	2.73333*	1.05389	.034	.1729
behaviour-based RS	complementary RS	-2.66667*	1.05389	.040	-5.2271
	control	.06667	1.05389	.998	-2.4938
control	complementary RS	-2.73333*	1.05389	.034	-5.2938
	behaviour-based RS	-.06667	1.05389	.998	-2.6271

The effect size ($\eta^2 = 0.172$) shows that about 17.2% of the changes in relationship well-being scores were due to the type of recommender system used, which is considered a large effect. Further analysis with a one-way ANOVA and post-hoc Tukey's HSD tests revealed significant differences in relationship change scores between the complementary RS group and the behaviour-based RS and control groups. The complementary RS group had a greater increase in scores (mean difference = +2.4) compared to the other two groups.

Sense of Meaning

As shown in table 23, post-questionnaire scores reveal a significant change. The complementary RS group increased substantially to 6.07, whereas the behaviour-based RS and control groups had more modest increases (5.53 and 4.73, respectively).

Table 23.

Descriptive Statistics

	group	Mean	Std. Deviation	N
pre_meaning	complementary RS	3.60	1.682	15
	behaviour-based RS	4.80	2.305	15
	control	4.87	2.326	15
	Total	4.42	2.158	45
post_meaning	complementary RS	6.07	1.335	15
	behaviour-based RS	5.53	1.995	15
	control	4.73	1.981	15
	Total	5.44	1.841	45

Table 24.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
time	Sphericity Assumed	23.511	1	23.511	6.051	.018
	Greenhouse-Geisser	23.511	1.000	23.511	6.051	.018
	Huynh-Feldt	23.511	1.000	23.511	6.051	.018
	Lower-bound	23.511	1.000	23.511	6.051	.018
time * group	Sphericity Assumed	26.289	2	13.144	3.383	.043
	Greenhouse-Geisser	26.289	2.000	13.144	3.383	.043
	Huynh-Feldt	26.289	2.000	13.144	3.383	.043
	Lower-bound	26.289	2.000	13.144	3.383	.043
Error(time)	Sphericity Assumed	163.200	42	3.886		
	Greenhouse-Geisser	163.200	42.000	3.886		
	Huynh-Feldt	163.200	42.000	3.886		
	Lower-bound	163.200	42.000	3.886		

The within-subject effects show a significant interaction between "time" and "group" ($p = .043$), indicating that score changes over time varied by group. An η^2 of 0.138 indicates a relatively large effect size which suggests that a significant portion of the variance in meaning change scores can be explained by the independent variable (recommendation system). One-way ANOVA and post-hoc Tukey's HSD tests further identified significant differences. Specifically, the complementary RS group showed a greater increase in meaning change scores (mean difference = 2.60) compared to the control group ($p = .037$). No significant differences were found between the complementary RS and behaviour-based RS.

Table 25.

ANOVA

meaning_change

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	52.578	2	26.289	3.383	.043
Within Groups	326.400	42	7.771		
Total	378.978	44			

Table 26.

Post Hoc Tests

Multiple Comparisons

Dependent Variable: meaning_change
Tukey HSD

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% ... Lower Bound
complementary RS	behaviour-based RS	1.73333	1.01793	.216	-.7397
	control	2.60000*	1.01793	.037	.1269
behaviour-based RS	complementary RS	-1.73333	1.01793	.216	-4.2064
	control	.86667	1.01793	.673	-1.6064
control	complementary RS	-2.60000*	1.01793	.037	-5.0731
	behaviour-based RS	-.86667	1.01793	.673	-3.3397

Achievement

Table 27 presents participants' achievement scores before and after the experiment, the complementary RS group had the lowest average pre-score (3.53), the behaviour-based RS group had the highest (4.87), and the control group was in between (4.60). Post-experiment, the complementary RS group's average score increased significantly to 5.73, while changes in the behaviour-based RS (4.47) and control groups (5.47) were more modest. A within-subjects ANOVA revealed a significant effect of time (p = .033), indicating a significant overall change in achievement scores from pre- to post-questionnaire. There was also a significant interaction effect between time and group (p = .040), suggesting differences in achievement score changes across groups.

Table 27

Descriptive Statistics

	group	Mean	Std. Deviation	N
pre_achievement	complementary RS	3.53	1.767	15
	behaviour-based RS	4.87	2.031	15
	control	4.60	2.063	15
	Total	4.33	2.000	45
post_achievement	complementary RS	5.73	1.624	15
	behaviour-based RS	4.47	2.167	15
	control	5.47	1.642	15
	Total	5.22	1.869	45

The effect size ($\eta^2 = 0.142$) means that 14.2% of the changes in achievement scores were due to the type of recommender system used, which is considered a large effect and indicates that the type of system had a substantial impact on how much participants felt a sense of achievement. Also, a between-subjects ANOVA showed a significant effect of group ($p = .040$), indicating differences in achievement score changes among the groups. Tukey's HSD post-hoc tests revealed that the complementary RS group had a significantly greater increase in achievement scores (mean difference = 2.60) compared to the behaviour-based RS group ($p = .031$), suggesting a more substantial positive influence from the complementary recommender system (see Table 30).

Table 28.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
time	Sphericity Assumed	17.778	1	17.778	4.884	.033
	Greenhouse-Geisser	17.778	1.000	17.778	4.884	.033
	Huynh-Feldt	17.778	1.000	17.778	4.884	.033
	Lower-bound	17.778	1.000	17.778	4.884	.033
time * group	Sphericity Assumed	25.356	2	12.678	3.483	.040
	Greenhouse-Geisser	25.356	2.000	12.678	3.483	.040
	Huynh-Feldt	25.356	2.000	12.678	3.483	.040
	Lower-bound	25.356	2.000	12.678	3.483	.040
Error(time)	Sphericity Assumed	152.867	42	3.640		
	Greenhouse-Geisser	152.867	42.000	3.640		
	Huynh-Feldt	152.867	42.000	3.640		
	Lower-bound	152.867	42.000	3.640		

Table 29.

Post Hoc Tests

Multiple Comparisons

Dependent Variable: achievement_change

Tukey HSD

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% ... Lower Bound
complementary RS	behaviour-based RS	2.60000*	.98518	.031	.2065
	control	1.33333	.98518	.374	-1.0602
behaviour-based RS	complementary RS	-2.60000*	.98518	.031	-4.9935
	control	-1.26667	.98518	.411	-3.6602
control	complementary RS	-1.33333	.98518	.374	-3.7268
	behaviour-based RS	1.26667	.98518	.411	-1.1268

DISCUSSION

We aimed to compare the impact of complementary and behaviour-based recommender systems on users' well-being. The results from the positive emotion analysis provide strong validation for our hypothesis. This finding is particularly interesting considering that the complementary recommendations specifically targeted participants with lower well-being scores in specific PERMA domains. Therefore, the complementary recommender system's ability to target specific well-being needs likely played a crucial role, as it can recommend movies that could potentially boost positive emotions, address underlying deficiencies, and contribute to a more significant emotional uplift compared to the behaviour-based recommendations. This new recommendation method can potentially impact users' daily lives by making them feel happier, more motivated, and less stressed, etc.

Our second objective was to analyse the impact on user engagement, hypothesising that behaviour-based algorithms would outperform. However, there was no differential impact on engagement. Despite the decreased engagement for the complementary group and a slight increase for the behaviour-based group, these were not statistically significant. Hence, we can conclude that we fail to validate the hypothesis and the behaviour-based recommender system did not lead to a statistically significant increase in user engagement compared to any other groups. This lack of statistical difference is important, as it may challenge the common belief that prioritising well-being can significantly reduce watch time or engagement on streaming platforms.

Next, we examined the impact on relationship satisfaction, hypothesising that the complementary system would improve relational well-being more. The results on relationship well-being supported our hypothesis. Although there was no significant overall change across all groups, there was a significant interaction effect between "time" and "group." Further analysis using ANOVA and Tukey's HSD revealed that, despite starting with a lower baseline, the complementary group showed significant improvement, validating our hypothesis.

Regarding sense of meaning, we hypothesised that the complementary system would positively impact it. While descriptive statistics suggested a notable difference in the mean sense of meaning scores between the complementary and behaviour-based recommender system groups, statistical analysis did not reveal a significant difference. However, since there's a significant difference between the control group and the CRS group, we can conclude that the complementary system is effective in enhancing users' sense of meaning. But, the absence of differences between the complementary and behaviour-based systems suggests their

Post Hoc Tests

Table 30.

Multiple Comparisons

Dependent Variable: achievement_change
Tukey HSD

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% ... Lower Bound
complementary RS	behaviour-based RS	2.60000*	.98518	.031	.2065
	control	1.33333	.98518	.374	-1.0602
behaviour-based RS	complementary RS	-2.60000*	.98518	.031	-4.9935
	control	-1.26667	.98518	.411	-3.6602
control	complementary RS	-1.33333	.98518	.374	-3.7268
	behaviour-based RS	1.26667	.98518	.411	-1.1268

comparable effectiveness in influencing users' sense of meaning, which weakens the ability to validate the hypothesis. Lastly, for sense of achievement, Tukey's HSD test revealed that the complementary group had a significantly greater increase in achievement scores compared to the behaviour-based group. This was aligned with our hypothesis. Interestingly, the behaviour-based group experienced a decrease in achievement scores, while the control group showed a slight increase, highlighting the potential benefits of stepping outside one's usual preferences.

CONCLUSION

Our findings offer valuable insights into how different recommender systems can influence users' well-being. The complementary recommender system led to a significant increase in positive emotions compared to the behaviour-based system. This suggests the system's ability to address specific well-being needs by recommending movies that evoke positive narratives or themes. There was no statistically significant difference in user engagement between the recommender system groups and the control group. The complementary recommender system led to a greater increase in relationship satisfaction compared to both the behaviour-based and control groups, which means exposure to positive and diverse relationship dynamics in movies can positively influence user perceptions of their own relationships. The complementary recommender system led to a significant increase in sense of meaning compared to the control group, partially validating the hypothesis. However, there was no significant difference between the complementary and behaviour-based groups. Lastly, the complementary recommender system led to a significantly greater increase in sense of achievement compared to the behaviour-based system, indicating that recommended movies potentially contributed to feelings of accomplishment, learning new things, or inspiring goal pursuit. Regarding implications, this research highlights the potential of recommender systems to go beyond user engagement and influence various aspects of well-being. By prioritising user needs and incorporating well-being considerations, recommender systems can be designed to not only keep users engaged but also contribute to their overall well-being. These findings inform developers, researchers, and platforms to consider well-being alongside user preferences when designing and implementing recommender systems as they can play a positive role in user experience and well-being in the digital age.

REFERENCES

Ansari, A., Essegaier, S. and Kohli, R. (2000). Internet Recommendation *Systems. Journal of Marketing Research*, 37(3), pp.363–375. doi:https://doi.org/10.1509/jmkr.37.3.363.18779.

Banker, S. and Khetani, S. (2019). Algorithm Overdependence: How the Use of Algorithmic Recommendation Systems Can Increase Risks to Consumer Well-Being. *Journal of Public Policy & Marketing*, 38(4), p.074391561985805. doi:https://doi.org/10.1177/0743915619858057.

B. Patel, P. Desai and U. Panchal, "Methods of recommender system: A review," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-4, doi: 10.1109/ICIIECS.2017.8275856.

Chatterjee, S. (2023). *Recommendation Systems: Ethical Challenges and the Regulatory Landscape*. [online] www.holisticai.com. Available at: https://www.holisticai.com/blog/recommendation-systems.

Chen, L., Wu, W. and He, L. (2013). How personality influences users' needs for recommendation diversity? doi:https://doi.org/10.1145/2468356.2468505.

Christiano, T. (2021). Algorithms, Manipulation, and Democracy. *Canadian Journal of Philosophy*, 52(1), pp.1–16. doi:https://doi.org/10.1017/can.2021.29.

Chu, S.L., Brown, S., Park, H. and Spornhauer, B. (2020). Towards Personalized Movie Selection for Wellness: Investigating Event-Inspired Movies. *International Journal of Human–Computer Interaction*, 36(16), pp.1514–1526. doi:https://doi.org/10.1080/10447318.2020.1768665.

CMA (2021). *Algorithms: how they can reduce competition and harm consumers Summary of responses to the consultation*.

Durrani, A. (2023). *The Average American Spends Over 13 Hours A Day Using Digital Media—Here's What They're Streaming*. [online] Forbes Home. Available at: https://www.forbes.com/home-improvement/internet/streaming-stats/.

eSafety Commissioner (2022). *Recommender systems and algorithms – position statement | eSafety Commissioner*. [online] eSafety Commissioner. Available at: https://www.esafety.gov.au/industry/tech-trends-and-challenges/recommender-systems-and-algorithms.

Fernández-Aguilar, L., Navarro-Bravo, B., Ricarte, J., Ros, L. and Latorre, J.M. (2019). How effective are films in inducing positive and negative emotional states? A meta-analysis. *PLOS ONE*, 14(11), p.e0225040. doi:https://doi.org/10.1371/journal.pone.0225040.

Ferwerda, B., Tkalcic, M. and Schedl, M. (2017). Personality Traits and Music Genres. *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization - UMAP '17*. [online] doi:https://doi.org/10.1145/3079628.3079693.

Gómez-Aguilar, A. (2020). Content Bubbles. *Advances in business strategy and competitive advantage book series*, p.339. doi:https://doi.org/10.4018/978-1-7998-3119-8.ch022.

Grobler, L. (2012). *An exploration of the use of positive psychology movies to enhance well–being*. [online] repository.nwu.ac.za. Available at: https://repository.nwu.ac.za/handle/10394/9051 [Accessed 24 May 2024].

Guo, L. (2018). Beyond the top-n: Algorithms that generate recommendations for self-actualization. In *Proceedings of the 12th acm conference on recommender systems* (pp. 573–577). ACM. doi:10.1145/3240323.3240330 (Guo, 2018)

Haddad, E. (2022). Escaping the echo chamber in movies recommender system. *Ndu.edu.lb*. [online] doi:http://ir.ndu.edu.lb/123456789/1584.

Haidt, J. and Allen, N. (2020). Scrutinizing the effects of digital technology on mental health. *Nature*, [online] 578(7794), pp.226–227. doi:https://doi.org/10.1038/d41586-020-00296-x.

Hajjar, S.T., (2018). Statistical analysis: Internal-consistency reliability and construct validity. *International Journal of Quantitative and Qualitative Research Methods*, *6*(1), pp.27-38.

Hamilton, J. (2023). *Watching movies could be good for your mental health*. [online] The Conversation. Available at: https://theconversation.com/watching-movies-could-be-good-for-your-mental-health-214325.

Harris, A.D., McGregor, J.C., Perencevich, E.N., Furuno, J.P., Zhu, J., Peterson, D.E. and Finkelstein, J. (2006). The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics. *Journal of the American Medical Informatics Association*, [online] 13(1), pp.16–23. doi:https://doi.org/10.1197/jamia.m1749.

Hemans, M. and Wiredu Ocansey, D.K. (2021). The Impact of Recommendation System Overuse on the Subjective Wellbeing of Internet Users. *International Journal of Scientific Research in Science, Engineering and Technology*, 8(1), pp.239–247. doi:https://doi.org/10.32628/ijsrset218156.

Henry, N., Pedersen, M., Williams, M., Martin, J. and Donkin, L. (2023). Reducing Echo Chamber Effects: An Allostatic Regulator for Recommendation Algorithms. *Research Square (Research Square)*. doi:https://doi.org/10.21203/rs.3.rs-3708350/v1.

Isinkaye, F.O., Folajimi, Y.O. and Ojokoh, B.A. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 16(3), pp.261–273. doi:https://doi.org/10.1016/j.eij.2015.06.005.

Karanicolas, P.J., Farrokhyar, F. and Bhandari, M. (2010). Practical Tips for Surgical research: blinding: who, what, when, why, how? *Canadian Journal of surgery. Journal Canadien De Chirurgie*, [online] 53(5), pp.345–348. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2947122/.

Krook, J. and Blockx, J. (2023). Recommender Systems, Autonomy and User Engagement. doi:https://doi.org/10.1145/3597512.3599712.

L. Heffner, C. (2022). *Doomscrolling: The Ultimate Negative Flow State*. [online] AllPsych. Available at: https://allpsych.com/doomscrolling-the-ultimate-negative-flow-state/.

Laerd Statistics (2018). *How to perform a Mixed ANOVA in SPSS Statistics | Laerd Statistics*. [online] Laerd.com. Available at: https://statistics.laerd.com/spss-tutorials/mixed-anova-using-spss-statistics.php.

Lampropoulos, G.K., Kazantzis, N. and Deane, F.P. (2004). Psychologists' Use of Motion Pictures in Clinical Practice. *Professional Psychology: Research and Practice*, [online] 35(5), pp.535–541. doi:https://doi.org/10.1037/0735-7028.35.5.535.

Lewczuk, K., Wójcik, A. and Gola, M. (2021). Increase in the Prevalence of Online Pornography Use: Objective Data Analysis from the Period Between 2004 and 2016 in Poland. *Archives of Sexual Behavior*, 51(2), pp.1157–1171. doi:https://doi.org/10.1007/s10508-021-02090-w.

Lex, E. and Schedl, M. (2022). Psychology-informed Recommender Systems Tutorial. *Sixteenth ACM Conference on Recommender Systems*. doi:https://doi.org/10.1145/3523227.3547375.

LIU, J., TANG, W., CHEN, G., LU, Y., FENG, C. and TU, X.M. (2016). Correlation and agreement: overview and clarification of competing concepts and measures. *Shanghai Archives of Psychiatry*, [online] 28(2), pp.115–120. doi:https://doi.org/10.11919/j.issn.1002-0829.216045.

Lorenz-Spreen, P., Mønsted, B.M., Hövel, P. and Lehmann, S. (2019). Accelerating dynamics of collective attention. *Nature Communications*, [online] 10(1), pp.1–9. doi:https://doi.org/10.1038/s41467-019-09311-w.

Lukoff, K. (2021). Addressing Present Bias in Movie Recommender Systems and Beyond. *IUI Workshops*.

Lyngs, U., Binns, R., Van Kleek, M. and Shadbolt, N. (2018). 'So, Tell Me What Users Want, What They Really, Really Want!' *Extended Abstracts of the 2018 CHI Conference on*

Human Factors in Computing Systems - CHI '18. [online] doi:https://doi.org/10.1145/3170427.3188397.

Main, P. (2023). *Self-Actualization*. [online] www.structural-learning.com. Available at: https://www.structural-learning.com/post/self-actualization#:~:text=Self%2Dactualization%20allows%20individuals%20to.

Mangot, A.G. and Murthy, V.S. (2017). Cinema: A multimodal and integrative medium for education and therapy. *Annals of Indian Psychiatry*, [online] 1(1), p.51. doi:https://doi.org/10.4103/aip.aip_13_17.

Mark, M.M. and Reichardt, C.S. (2001). *Internal Validity Threat - an overview | ScienceDirect Topics*. [online] www.sciencedirect.com. Available at: https://www.sciencedirect.com/topics/computer-science/internal-validity-threat#:~:text=.

Marples, M. (2021). *Doomscrolling can steal hours of your time -- here's how to take it back*. [online] CNN. Available at: https://edition.cnn.com/2021/02/26/health/doomscrolling-prevention-tips-wellness/index.html [Accessed 11 Feb. 2024].

Martinez, S. (2021). *Streaming Service Algorithms are Biased, Directly Affecting Content Development*. [online] AMT Lab @ CMU. Available at: https://amt-lab.org/blog/2021/11/streaming-service-algorithms-are-biased-and-directly-affect-content-development.

Mernin, A. (2020). *Can a VR headset detect emotions? - Health Tech World*. [online] www.htworld.co.uk. Available at: https://www.htworld.co.uk/news/can-a-vr-headset-detect-emotions/ [Accessed 20 Feb. 2024].

Metzler, H., Pellert, M. and Garcia, D. (2022). *Using social media data to capture emotions before and during COVID-19*. [online] worldhappiness.report. Available at: https://worldhappiness.report/ed/2022/using-social-media-data-to-capture-emotions-before-and-during-covid-19/.

Michael D. Ekstrand and Martijn C. Willemsen. (2016). Behaviorism is Not Enough: Better Recommendations through Listening to Users. In Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16). Association for Computing Machinery, New York, NY, USA, 221–224. https://doi.org/10.1145/2959100.2959179

Mintel (2023). *Video Streaming Services - UK - 2023: Overview*. [online] reports.mintel.com. Available at: https://reports.mintel.com/display/1200653/?fromSearch=%3Ffreetext%3DVOD%26resultPosition%3D2 [Accessed 27 Feb. 2024].

Mourya, A. (2023). *A Survey on Movie Recommendation System*. [online] Medium. Available at: https://medium.com/@akshaymouryaart/a-survey-on-movie-recommendation-system-d9610777f8e5 [Accessed 3 Mar. 2024].

Murrar, S. and Brauer, M. (2018). The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation Mixed Model Analysis of Variance Title: The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation Chapter Title: 'Mixed Model Analysis of Variance'. *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. [online] doi:https://doi.org/10.4135/9781506326139.n436.

Netflix (2022). *How Netflix's Recommendations System Works*. [online] Help Center. Available at: https://help.netflix.com/en/node/100639.

Newsom , J.T. (2023). *Newsom: Univariate Statistics*. [online] web.pdx.edu. Available at: https://web.pdx.edu/~newsomj/uvclass/ [Accessed 13 May 2024].

Ng, J. (2024). *Content-based Recommender Using Natural Language Processing (NLP)*. [online] KDnuggets. Available at: https://www.kdnuggets.com/2019/11/content-based-recommender-using-natural-language-processing-nlp.html.

Nguyen, T.T., Hui, P.-M., Harper, F.M., Terveen, L. and Konstan, J.A. (2014). Exploring the filter bubble. *Proceedings of the 23rd international conference on World wide web - WWW '14*. [online] doi:https://doi.org/10.1145/2566486.2568012.

Nha, V.T.T. (2021). Understanding validity and reliability from qualitative and quantitative research traditions. *VNU Journal of Foreign Studies*, [online] 37(3). doi:https://doi.org/10.25073/2525-2445/vnufs.4672.

Niemiec, R.M. (2020). Character strengths cinematherapy: Using movies to inspire change, meaning, and cinematic elevation. *Journal of Clinical Psychology*, 76(8), pp.1447–1462. doi:https://doi.org/10.1002/jclp.22997.

Niemiec, R.M. and Wedding, D. (2014). *Positive psychology at the movies : using films to build character strengths and well-being*. Toronto, Ontario: Hogrefe Publishing.

Noordeh, E., Levin, R., Jiang, R. and Shadmany, H. (2020). Echo Chambers in Collaborative Filtering Based Recommendation Systems. *arXiv (Cornell University)*. doi:https://doi.org/10.48550/arxiv.2011.03890.

NVIDIA (2024). *What is a Recommendation System?* [online] NVIDIA Data Science Glossary. Available at: https://www.nvidia.com/en-us/glossary/recommendation-system/.

Page, S.J. and Persch, A.C. (2013). Recruitment, Retention, and Blinding in Clinical Trials. *American Journal of Occupational Therapy*, [online] 67(2), pp.154–161. doi:https://doi.org/10.5014/ajot.2013.006197.

Pannell, R. (2023). *The Importance of Identifying the Right Sample Size for Business Improvement - LeanScape*. [online] Leanscape. Available at: https://leanscape.io/the-importance-of-identifying-the-right-sample-size-for-business-improvement/#:~:text=Why%20is%2030%20the%20minimum.

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.

Pattrn.AI (2023). *Evaluating recommender systems in relation to illegal and harmful content*. [online] Ofcom. Available at: https://www.ofcom.org.uk/research-and-data/online-research/evaluating-recommender-systems-in-relation-to-the-dissemination-of-illegal-and-harmful-content-in-the-uk [Accessed 14 Mar. 2024].

Polignano, M., Narducci, F., Gemmis, M. and Semeraro, G. (2020). Towards Emotion-aware Recommender Systems: an Affective Coherence Model based on Emotion-driven Behaviors. *Expert Systems with Applications*, p.114382. doi:https://doi.org/10.1016/j.eswa.2020.114382.

Poon, C. (2022). *Viewers' Dilemma*. [online] Medium. Available at: https://medium.com/@cowin/the-feeling-you-have-when-scrolling-through-netflix-or-another-other-streaming-service-trying-to-1279899d38af.

Popa, D., Nechita, F., Liu, Y. and Wei Lee Chin, S. (2021). Linking Positive Psychology and Intercultural Competence by Movies: Evidence From Brunei and Romania. *Frontiers in Psychology*, [online] 12, p.750904. doi:https://doi.org/10.3389/fpsyg.2021.750904.

Reichardt, C.S. and Little, T.D. (2019). *Quasi-experimentation : a guide to design and analysis*. New York: The Guilford Press.

Reisach, U. (2020). The Responsibility of Social Media in Times of Societal and Political Manipulation. *European Journal of Operational Research*, [online] 291(3). doi:https://doi.org/10.1016/j.ejor.2020.09.020.

Rodriguez, A. (2018). *People still rely on other people more than algorithms to find TV shows to watch*. [online] Quartz. Available at: https://qz.com/1219894/study-people-beat-algorithms-when-it-comes-to-recommending-tv-shows [Accessed 22 Feb. 2024].

Saurwein, F. and Spencer-Smith, C. (2021). Automated Trouble: The Role of Algorithmic Selection in Harms on Social Media Platforms. *Media and Communication*, 9(4), pp.222–233. doi:https://doi.org/10.17645/mac.v9i4.4062.

Seariac, H. (2024). *As Zuckerberg apologizes in hearing, Utah suit alleges Meta knew its apps harm children*. [online] Deseret News. Available at: https://www.deseret.com/2024/2/1/24058701/meta-mark-zuckerberg-ceo-utah-lawsuit [Accessed 14 Feb. 2024].

Smith, D. (2023). *Eight Reasons Why Watching Movies Is Beneficial For Mental Health*. [online] The Art of Healthy Living. Available at: https://artofhealthyliving.com/eight-reasons-why-watching-movies-is-beneficial-for-mental-health/ [Accessed 29 Feb. 2024].

Thomas, L. (2022). *Quasi-Experimental Design | Definition, Types & Examples*. [online] Scribbr. Available at: https://www.scribbr.co.uk/research-methods/quasi-experimental/.

Thorat, P. B., Goundar, R. M., & Barve, S. (2015). Survey on collaborative filtering content-based filtering and hybrid recommendation system. International Journal of Computer Applications, 110, 31-36.

UCLA (2021). *What does Cronbach's alpha mean? | SPSS FAQ*. [online] stats.oarc.ucla.edu. Available at: https://stats.oarc.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean/#:~:text=Cronbach.

Victoria McNally (2021). *Who do people trust the most when it comes to movie and TV recommendations? Hint: It's not critics*. [online] digitalhub US. Available at: https://swnsdigital.com/us/2021/04/who-do-people-trust-the-most-when-it-comes-to-movie-and-tv-recommendations-hint-its-not-critics/.

Vidiyala, R. (2020). *How to Build a Movie Recommendation System?* [online] Medium. Available at: https://towardsdatascience.com/how-to-build-a-movie-recommendation-system-67e321339109.