A Work Project, presented as part of the requirements for the Award of a Master's degree in Management from the Nova School of Business and Economics.

# SURPRISE ME! A QUANTITATIVE STUDY OF USER SATISFACTION WITH SERENDIPITOUS RECOMMENDATIONS ACROSS MAJOR DIGITAL PLATFORMS.

ANYA LEAH BERCHTOLD

Work project carried out under the supervision of:

Professor João Castro

17/05/2025

**Abstract**

While personalization has advanced, users often report unmet expectations for discovery in recommendations across digital platforms. This thesis addresses the research gap and explores how users perceive serendipitous content discovery – or 'nice surprises' – across services like Netflix, Spotify, Amazon, LinkedIn, and YouTube. The quantitative study revealed a significant gap between expected and experienced serendipity, strongly linked to lower satisfaction. Curiosity, trust, and platform context further shaped user evaluations. The findings inform a multi-phase practical guide to help platforms better integrate serendipity without compromising personalization, enhancing user experience, engagement, and long-term retention.

**Keywords**

recommender system, digital platforms, serendipity, user satisfaction, user acceptance, expectancy disconfirmation, Netflix, YouTube, Spotify, Amazon, LinkedIn

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

## Table of Contents

# 1. Introduction

As digital platforms present users with an unprecedented array of choices, the algorithmic recommender system (RS) has emerged as an essential tool for filtering information and guiding user decisions. These machine-learning systems promise to ease the information overload by delivering content tailored to individual preferences, often boosting business performance through increased engagement and conversion rates (Amatriain and Basilico 2015; Boudet et al. 2020). However, while personalization is central to digital strategy, its dominance raises concerns about overfitting and recommendation redundancy. Users frequently report repetitive recommendations and frustration with algorithmic monotony, as highlighted by Reddit threads and recent articles such as “How To Stop Spotify Feeding You The Same Old Songs” (Collins 2020), “Retraining my YouTube algorithm saved my sanity” (Hachey 2025), “Here’s How I Reset My Netflix Recommendations When They Aren’t Good Anymore” (Wanjala 2024), and “Amazon using AI to fix its broken Prime Video algorithm” (Donaldson 2024). These frustrations persist despite the theoretical potential of RSs to foster discovery and surprise, and the consistent claim from platforms that they promote these aspects. Serendipity, or the delivery of relevant yet unexpected content, has gained attention as a key dimension against this overspecialization problem, yet its practical realization appears limited. If users actively seek ways to “retrain” their algorithms, it suggests that platforms may not deliver the level of serendipitous discovery that users expect, compromising user satisfaction.

Previous academic research has focused mainly on defining serendipity in RSs (e.g., Kotkov et al. 2024; Stitini et al. 2023), various methods to implement it in the RSs (e.g., Kim et al. 2017; Ziarani and Ravanmehr 2021a), and exploring its potential to enhance user satisfaction (e.g., Kotkov et al. 2018; Ping, Li, and Zhu 2024). Studies have investigated how factors like novelty, relevance, and diversity contribute to serendipitous experiences, and some have begun to address individual differences in user preferences. However, there is still limited understanding of how much users want it and whether they are satisfied with the current level. The expected findings aim to fill this gap and help outline the user need, laying the groundwork for designing RSs in ways users value.

The research question that drives this thesis is: To what extent are users satisfied with the level of serendipitous recommendations in modern digital platforms? The following objectives are also considered: How do users anticipate serendipity in their interactions with digital platforms? How do users perceive the degree to which platforms deliver serendipitous experiences? What factors influence user satisfaction with serendipitous recommendations?

This thesis has six sections. The literature review (2) introduces key concepts surrounding serendipity in RSs and user satisfaction, followed by the quantitative methodology section (3). The statistical results are then presented (4), which are discussed in terms of key findings, implications, and limitations (5), concluding with a summary (6).

# 2. Literature Review

The literature for this thesis consists of articles and journal publications retrieved from various academic databases, such as ACM Digital Library, Emerald, ScienceDirect, and Springer. Papers are screened to ensure that only those published from 2000 onwards and in English are included.

## 2.1 Serendipity in Recommender Systems

### 2.1.1 Overview of Recommendation Algorithms

A recommender system (RS) is a software tool that suggests items to users to help them make informed decisions when faced with multiple choices and information overload (Javed and Ashraf 2023; Ricci, Rokach, and Shapira 2011). Over time, they have evolved into strategic tools for business to drive key metrics like engagement, retention, and revenue (Amatriain and Basilico 2015; Gorgoglione, Panniello, and Tuzhilin 2019).

Research categorizes the algorithms within RSs into three main types. *Collaborative filtering* (CF) makes item suggestions based on similarities between users or items from historical interactions; this is the most widely used RS technique (Ping, Li, and Zhu 2024). *Content-based filtering* (CBF) uses items the user has previously engaged with and rated to make suggestions of items with similar characteristics. *Hybrid models* combine both to mitigate their limitations, such as reduced exposure to diversity or struggle with data sparsity (Afoudi, Lazaar, and Al Achhab 2021), thus improving accuracy and effectiveness (Ping, Li, and Zhu 2024). RS algorithms can exhibit exploitative (reinforcing known preferences) or explorative (introducing novel content) behavior, with multi-armed bandit models balancing both by testing multiple options, or 'arms', and then exploiting the best-performing option (McInerney et al. 2018).

Overspecialization remains a significant challenge in RSs, particularly those relying heavily on algorithms reinforcing past behavior, leading to "filter bubbles" and reduced user engagement (Kim et al. 2017; Stitini et al. 2023). The term, coined by Pariser (2011, cited in Arguedas et al. 2022), describes how personalization can limit exposure to diverse viewpoints, contributing to social fragmentation. Scholars remain divided on the severity of this effect - some support Pariser's claim that filter bubbles reinforce biases (e.g., Krook 2025; Reviglio 2019), while others argue RSs can still promote diversity through random discovery (e.g., Arguedas et al. 2022; McInerney et al. 2018). Fleder, Hosanagar, and Buja (2011) and Hazrati and Ricci (2024) suggest that RSs converge users around popular content, even when shown niche recommendations. This suggests that user agency and self-selection may be more responsible for perspective narrowing than the algorithms themselves. Still, modern research has increasingly focused on methods to mitigate overspecialization and promote a more balanced user experience.

#### 2.1.2 Evaluating Recommender Systems

RSs have traditionally been evaluated using accuracy metrics; however, an overemphasis on accuracy can lead to redundant recommendations (Konstan and Riedl 2012; Ping, Li, and Zhu

2024; Ziarani and Ravanmehr 2021a). As a result, the focus has shifted to "beyond-accuracy" metrics of novelty, diversity, and serendipity. Among these, serendipity remains particularly difficult to evaluate due to its subjective and emotional nature, with no standardized metrics or formulas (Kotkov et al. 2024; Stitini et al. 2023; Ziarani and Ravanmehr 2021b). User feedback-based evaluations are considered the most reliable, though they demand more effort and implementation costs (Kotkov et al. 2018; Nalis et al. 2024).

Reflecting this shift, modern RS evaluations increasingly incorporate A/B testing (i.e., split testing two versions) and user feedback, acknowledging that offline accuracy does not fully capture engagement or satisfaction (Konstan and Riedl 2012; Ricci, Rokach, and Shapira 2011). Recent studies emphasize the use of user-centric frameworks such as the TAM (Technology Acceptance Model), UTAUT (Unified Theory of Acceptance and Use of Technology), and ResQue model (Recommender System's Quality of User Experience) tailored for RSs (Pu, Chen, and Hu 2013). These models capture user perceptions and satisfaction directly, connecting algorithmic performance and actual user experience.

#### 2.1.3 Serendipity as an Objective

To counter this algorithmic trap of overspecialization, users need something unexpected - this is where serendipity comes in. Serendipity is "the faculty or phenomenon of finding valuable or agreeable things not sought for" (Merriam-Webster.com Dictionary, n.d.), but RS research lacks a universal definition, leading to varied interpretations. This paper combines them and defines serendipity in RSs as *the enhancement of discovery through relevant, diverse, unexpected, and novel recommendations, ensuring positively surprising and useful suggestions* (Nalis et al. 2024; Ping, Li, and Zhu 2024; Ziarani and Ravanmehr 2021b).

Although the perception of serendipity is highly subjective and complicated to manipulate artificially, progress has been made in research. Beyond a natural exploration effect from reinforcement learning in certain algorithms, several methods aim to improve serendipity-based recommendations while mitigating filter bubbles: multi-armed bandit algorithms balancing exploration and exploitation (McInerney et al. 2018), context-aware approaches (Armentano et al. 2014; Konstan and Riedl 2012; Stitini et al. 2023), and deep learning and serendipity-focused re-ranking (Stitini et al. 2023; Ziarani and Ravanmehr 2021a). Additionally, strategies such as controlled randomness or diversity and hybrid human-AI curation can promote unexpected yet relevant content (Kim et al. 2017; Malekzadeh Hamedani and Kaedi 2019; Nagulendra and Vassileva 2014). Digital design can also facilitate serendipity through diverse content exposure, flexible navigation, and user-driven customization (Smets et al. 2022). Platforms like Spotify offering tools for user control (e.g., "Remove from taste profile"), curated RSs (Kim et al. 2017), and transparency-enhancing visualization techniques (Nagulendra and Vassileva 2014; Smets et al. 2022) emphasize that awareness alone is insufficient to counteract algorithmic loops - genuine discovery requires user agency and participation.

### 2.1.4 Applications in Digital Platforms

Data on how digital platforms integrate serendipity is limited; based on how serendipity can be facilitated, the following outlines the potential techniques each uses to foster user discovery. Platforms are chosen based on literature support and popularity within each of their industries.

Table 1. Applications of RSs in Digital Platforms

| Platform | RS Type | Description |
|---|---|---|
| **Netflix** | Hybrid | • Initially relied on explicit user ratings during DVD era; evolved with streaming and insights from the Netflix Prize (2006-2009), which exposed the limits of accuracy metrics (Amatriain and Basilico 2015; Gomez-Uribe and Hunt 2015; Gorgoglione, Panniello, and Tuzhilin 2019).<br>• Now uses a hybrid RS combining CF (based on watch history) and CBF (matching item characteristics to user preferences), resulting in a highly personalized interface (Kim et al. 2017). |
| **YouTube** | CF and Deep Learning | • RS drives ~60% of homepage clicks and employs a two-stage process (Covington, Adams, and Sargin 2016; Davidson et al. 2010):<br>1. CF first identifies candidate videos based on user behavior (watching, searching, liking, subscriptions).<br>2. Deep learning model then ranks candidates by predicted engagement, factoring in video quality, relevance, and diversity. |

| | | |
|---|---|---|
| | | • Recommendations are shown on homepage/browse pages with explanation and user controls. |
| **Spotify** | Hybrid | • Initially relied on CF but quickly integrated acoustic, contextual, user-interaction, and external data to refine their recommendations (Freeman, Gibbs, and Nansen 2022; Gorgoglione, Panniello, and Tuzhilin 2019).<br>• Employs a mix of AI and machine learning models to power features like Discover Weekly, Release Radar, and Daily Mix (Mudaliyar 2024).<br>• Despite the success of these features, users report overspecialization fatigue (Wasco and Read 2024). |
| **Amazon** | CF and Deep Learning | • Pioneered item-to-item CF in 1998, focusing on product similarities rather than user comparisons (Linden, Smith, and York 2003).<br>• This model increases scalability and efficiency, using co-purchase data and browsing/cart history for recommendations (Hardesty 2019; Kim et al. 2017; Linden, Smith, and York 2003).<br>• Prime Video evolved from CF and matrix factorization (user-item matrix) to deep learning post-2014 for improved prediction (Hardesty 2019). |
| **LinkedIn** | Hybrid | • Uses navigational datasets ("browsemaps") to track browsing patterns for job, connection, and content recommendations (Wu et al. 2014).<br>• Applies item-based CF for features like "People who viewed this profile also viewed" and "Jobs you may be interested in" (Gorgoglione, Panniello, and Tuzhilin 2019).<br>• CF is further enhanced with CBF factors such as industry, location, and experience (Wu et al. 2014), to balance personalization, engagement, and monetization. |

## 2.2 User Satisfaction

This section highlights some of the key researched influences on user acceptance, perception, engagement, behavior intentions, and satisfaction regarding RSs and serendipity. Based on these factors and the previous RS literature, hypotheses (H1-H9) are outlined in this section.

### 2.2.1 Expectancy Disconfirmation Theory

Finding the balance between personalization and discovery is critical for user satisfaction with RSs. While novelty and surprise can enhance engagement, excessive randomness or poorly aligned unexpectedness may reduce satisfaction (Chen et al. 2019; Kotkov et al. 2018; Ping, Li, and Zhu 2024). Serendipity-driven recommendations, which are not just novel or relevant but also pleasantly unexpected, have been shown to support preference broadening, increased engagement, and greater behavioral intent (Chen et al. 2019; Javed and Ashraf 2023; Kotkov et al. 2018). However, their positive effect on satisfaction depends on alignment with relevance, user expectations, and context, making precise tailoring essential in RS design.

Thus, expectancy disconfirmation theory (EDT, Oliver 1977, cited in Li, Choi, and Kim (2020) and Martínez-López et al. (2010)) offers a valuable lens for analyzing RS and serendipity satisfaction. Rooted in consumer satisfaction theory, it posits that satisfaction arises from comparing expected and perceived performance. If the performance meets or exceeds user expectations, satisfaction increases; if it underdelivers, this results in dissatisfaction. This model complements core constructs of acceptance models, especially the perceived performance variable 'perceived usefulness' (TAM) or 'performance expectancy' (UTAUT), and expectations are shaped by various inputs, including personality traits, social norms, and prior experiences (Olsson 2014). In their study, Li, Choi, and Kim (2020) suggest that high accuracy for an RS naturally leads to satisfaction by matching relevance (accuracy) expectations; however, over time, they found the overspecialization problem creates dissatisfaction, thus requiring diversity as a driver to increase behavior intent.

Studies such as Lavie et al. (2010, cited in Zanker et al. (2010)) and Lyngs et al. (2018) show that actual interactions often diverge from stated preferences, highlighting the need for RSs to adapt to dynamic user behavior. Kotkov et al. (2024) call attention to a persistent gap between what users expect from RSs and what they experience, framing this as a key challenge in RS evaluation. This study builds on that insight to examine whether popular platforms used for discovery successfully meet (or miss) user expectations and how this discrepancy shapes their satisfaction with serendipitous recommendations.

### 2.2.2 Influences on User Perceptions and Satisfaction

**Perceived performance**

Research on how users perceive and interact with RSs often draws from acceptance models, particularly the TAM and the UTAUT. A central concept in both models is the perceived performance ('perceived usefulness' or 'performance expectancy'), or the belief that using the system improves task performance. This performance perception has been shown to strongly predict behavioral intentions, such as the likelihood of returning to the platform or using recommended items (Armentano, Christensen, and Schiaffino 2015; Pu, Chen, and Hu 2013; Wang et al. 2012). Building on TAM, the ResQue model (Recommender System's Quality of User Experience) by Pu, Chen, and Hu (2013) adapts these principles specifically for RSs by incorporating specific factors such as recommendation quality, transparency, control, and trust. Like TAM, ResQue also finds that 'perceived usefulness' significantly influences behavioral intentions toward RSs. As previously mentioned, this construct plays a key role in the EDT.

*H1: Users' expectations of serendipitous discovery are significantly higher than their actual experienced (perceived) serendipity.*

*H2: Higher difference between expectation and experienced serendipity is associated with lower satisfaction with serendipity (serendipitous recommendations).*

**Control**

Having a sense of control over recommendations plays an important role in users' satisfaction, as they feel like they have kept some autonomy. Research shows that users who feel they can adjust or influence recommendations report greater trust and satisfaction with the system (Nagulendra and Vassileva 2014; Pu, Chen, and Hu 2013). However, too much control and excessive customization options can become overwhelming and counterproductive, leading to frustration instead of empowerment (Chen et al. 2013; Harambam et al. 2019; Pu, Chen, and Hu 2013). Freeman, Gibbs, and Nansen (2022) further observed that many users try to manage their recommendations (e.g., by creating multiple profiles or manually curating suggestions), often in response to distrust in the system's intentions or a decline in recommendation quality.

*H3: Higher difference between expectation and experienced control is associated with lower satisfaction with serendipity.*

## Contexts

Not all digital platforms are designed with the same content strategies or goals, which influence how their RSs are experienced. McCay-Peet, Toms, and Kelloway (2014; 2015) explore how the characteristics of digital environments shape the perception of serendipity, identifying several contextual factors. Building on those studies, Lutz, Hoffmann, and Meckel (2017) discovered that users find more satisfaction from serendipitous experiences on social media than in e-commerce, suggesting that the same recommendation features are perceived differently depending on the platform. This could be due to the frequency of social media use, which allows for more data collection and testing than e-commerce. In contrast, Javed and Ashraf (2023) found in their study on e-commerce and retail serendipitous recommendations did increase general engagement and satisfaction. Variations in both interface design (Harambam et al. 2019; Nalis et al. 2024; Smets et al. 2022) and underlying business goals (Gorgoglione, Panniello, and Tuzhilin 2019; Kim et al. 2017) further reinforce these contextual differences. In this light, this study investigates whether platform differences translate into distinct user experiences of serendipity and satisfaction with it.

*H4: There are significant differences of experienced serendipity between platforms.*

*H5: There are significant differences of satisfaction with serendipity between platforms.*

## Trust

Trust is a foundational component of user engagement with digital technologies and plays a similarly critical role in shaping how users perceive and adopt recommendations. In their user-centric RS evaluation framework, Pu, Chen, and Hu (2013) found that trust significantly influences purchase intention and overall usage. Other studies confirm this pattern: Martínez-López et al. (2010) identified trust as a key factor of user attitude toward RSs, while Wang et al. (2012) and Ricci, Rokach, and Shapira (2011) emphasized that trust directly enhances behavioral intention and user satisfaction. Serendipitous recommendations may increase engagement, but only if users perceive the system as trustworthy (Nalis et al. 2024). This could also translate into expectations, as users may expect more serendipity if they trust the platform to know them well. These findings taken together highlight trust as a recurring predictor of recommendation adoption, satisfaction, and perceived value.

*H6: Higher trust in the RS is positively associated with expectations of serendipity.*

*H7: Higher trust in the RS is positively associated with satisfaction with serendipity.*

**Curiosity**

User traits can play a significant role in how RS outputs are interpreted, and curiosity stands out as especially relevant in the context of serendipity. Research shows that highly curious individuals are more likely to embrace serendipitous recommendations, while users with narrower interests prefer familiar content (Ping, Li, and Zhu 2024; Stitini et al. 2023). Chen et al. (2019) found that curiosity fosters openness to surprise and novelty, which in turn makes unexpected recommendations more positively received. Meanwhile, McCay-Peet, Toms, and Kelloway's (2015) findings contradict this, stating that openness to experience does not have a relationship with the perception of serendipity. Reviglio (2019) supports the former, that serendipity emerges from both system design and a user's exploratory mindset. These findings suggest that curiosity may explain how open to (expectations) and how well users respond to (satisfaction) serendipitous content, making it a key individual-level influence in RS experience.

*H8: Higher curiosity is positively associated with expectations of serendipity.*

*H9: Higher curiosity is positively associated with satisfaction with serendipity.*

Beyond the above influences, studies also highlight that user engagement with RSs is influenced by other factors not included in this study, for example, cognitive biases.

## 2.3 Research Gap and Contribution

Although current research has explored methods for integrating serendipity into RSs, the role of technology acceptance in RS adoption, and the impact of serendipity on user satisfaction, a critical gap remains in understanding what users expect and whether they are truly satisfied with the serendipitous experiences offered by today's platforms. No study has systematically assessed whether users believe that popular digital platforms provide an adequate balance of personalization and unexpected discovery. This study addresses that gap by evaluating user expectations, perceptions, and satisfaction with serendipity, building on the expectancy-experience misalignment identified by Kotkov et al. (2024).

# 3. Methodology

This study adopts a quantitative approach (following Nenty (2009)) to collect structured, comparable data across a diverse participant group, enabling statistical analysis of relationships between expectations, perception, satisfaction, and user traits.

## 3.1 Research Design

A cross-platform user survey was developed using Qualtrics to test the proposed hypotheses, which was selected for its advanced logic and data export options. An initial description informed participants of the study's purpose and the option to enter a gift card raffle as appreciation for their time. Study participation required providing informed consent by clicking forward.

Participants first selected which platforms they use at least once per month (options randomized to minimize ordering effects), after which only answer options for those chosen were shown. Usage frequency and search behavior were collected to contextualize serendipity responses by linking them to engagement patterns, and platform awareness regarding data collection for personalized recommendations was also assessed. The core part of the survey used key constructs from previously validated measurements. Serendipity expectations and perceived performance (experiences) were adapted from Lutz, Hoffmann, and Meckel (2017), while control over personalization was assessed using an item from Pu, Chen, and Hu's ResQue model (2013). Satisfaction items integrated novelty, diversity, and unexpectedness elements, also based on ResQue. RS trust was measured using two ResQue items and curiosity using three statements from the CEI-II scale, following Chen et al. (2019). All core items used 5-point Likert scales.

An author-invented item directly addressed the perceived balance between repetition and discovery, and an open-ended text field allowed for any comments from the respondents. Demographic questions followed (nationality, age, and gender), but names were not collected for privacy. The final item thanked participants and included a field for contact information to enter the raffle. The full table of questions and their sources is in Appendix B Table 2.

## 3.2 Data Collection

A three-day pilot study was conducted before publishing the full survey. Due to feedback about imprecise wording and it being too lengthy, it was adjusted to enhance clarity and streamline the question format. A two-week data collection period started on April 2nd, 2025, and the survey was distributed through the author's network, social media, university group chats, and survey platforms. This mixed distribution helped recruit participants with diverse nationalities, age groups, and platform usage experiences. In total, 212 responses were collected.

## 3.3 Data Pre-Processing

Four responses were removed as their answers indicated a rushed click-through without consideration. The remaining 208 valid responses were cleaned in Excel of unnecessary information and checked for standardization. Per user, response averages and expectation-experience gaps were then calculated on the platform and aggregate levels (average across platforms), in preparation for data analysis.

## 3.4 Statistical Procedure

To test the hypotheses, a series of inferential statistical analyses were conducted using Jamovi software. The significance level for all tests was set at the common level of alpha (a) = 0.05.

Before performing the analyses, relevant assumptions were checked, including normality, to ensure correct measure usage and the robustness of the results. Shapiro-Wilk showed non-normality for all variables, so all analyses considered this.

# 4. Analysis and Results

## 4.1 Reliability and Validity

Cronbach's alpha was calculated for all multi-item variables across platforms to assess internal consistency, with the results shown in Appendix B Table 3. Although some coefficients fell slightly below the conventional 0.70 threshold, values between 0.60 and 0.80 are still considered acceptable (Janssens et al. 2008). Thus, the reliability results were considered acceptable for this quantitative analysis.

## 4.2 Descriptive Statistics

The sample was predominantly young (58.2% aged 18–24) and female (63.5%). Most participants were from Europe (54.8%), followed by North America (19.7%) and Asia (15.4%). YouTube was the most widely used platform (86.1%), followed by Spotify (67.3%), Netflix (64.9%), LinkedIn (55.3%), and Amazon (42.3%). All statistics are in Appendix B Table 4.

Platform subsamples revealed minor demographic and behavioral variations (Appendix B Table 5). The 18–24 age group dominated across platforms, highest among Spotify users (65.7%) and lowest for Amazon (53.4%), which also had the largest share of users above 35. Gender distribution was generally consistent, though Amazon showed a notably higher proportion of female users (68.2%) and the lowest male representation (28.4%). Non-binary participants (n=4) were present mainly among Netflix and YouTube users.

Spotify and YouTube had the highest daily usage rates (65.7% and 53.6%, respectively), compared to just 3.4% for Amazon. Exploration patterns varied: Spotify and Amazon users tended toward goal-directed searches, while users on Netflix, YouTube, and LinkedIn more often reported open or vague browsing, indicating greater potential for serendipity. Awareness of data-driven personalization was generally high across platforms, exceeding 80%, except for LinkedIn, where only 63.5% of users reported being aware.

On the 5-point scale from "more repetitive" to "more discovery," the perceived balance of recommendations across platforms was generally centered around the midpoint "balanced." Spotify (mean = 3.23), YouTube (3.17), and Netflix (3.00) leaned slightly toward discovery, while Amazon (2.95) and LinkedIn (2.97) skewed slightly toward repetitive recommendations. The open-ended responses (Appendix B Table 6) revealed frustrations with repetitive recommendations, lack of trust and transparency, and poor content quality. Respondents wanted platforms to balance personalization with discovery better and suggested features like recommendation vs. search labels, filtering by quality, or a toggling reset.

## 4.3 Hypothesis Testing

This section highlights the key results and outcomes from the statistical hypothesis testing, taking into account the observed non-normality. All test results are in Appendix B, Tables 7 – 15.

**H1 and H2:** A Wilcoxon signed-rank test revealed a statistically significant difference between expectation and experience for all platforms (each $p < .05$) and in the aggregate model ($p < .001$). These results support H1 and suggest that digital platforms tend to underdeliver on serendipity expectations. These expectation-experience gaps then showed a significant negative association with satisfaction across all platforms ($p < .05$), with the strongest effects observed for Spotify and Netflix and weakest for YouTube. The aggregate model also confirmed this relationship ($B = –0.233$, $p < .001$), affirming H2 and that unmet expectations of serendipity reduce user satisfaction.

**H3:** To examine whether discrepancies between expected and experienced control over the user's taste profile predicted lower satisfaction with serendipitous recommendations, linear regressions were run. The results revealed that the gap in perceived control over taste profile negatively predicted satisfaction only on Spotify ($B = –0.196$, $p = .002$), with a marginal effect in the aggregate model ($p = .051$). No significant effects were found for other platforms. H3 is, therefore, only partially supported and driven by Spotify.

**H4 and H5:** A non-parametric repeated measures ANOVA (Friedman test) revealed significant differences in perceived serendipity across platforms ($p = .004$). Post-hoc Dubin-Conover pairwise comparisons revealed that LinkedIn and Amazon were rated significantly lower than Netflix, YouTube, and Spotify. These results support H4 and indicate platform-dependent variation in perceived serendipitous experience. Another Friedman test confirmed significant differences in satisfaction with serendipitous recommendations across platforms ($p = .002$). Post-hoc pairwise comparisons showed that YouTube and Spotify outperformed (higher satisfaction) Amazon and LinkedIn, with no significant differences between YouTube, Spotify, and Netflix. H5 is thus supported.

**H6 and H7**: A series of linear regressions show a statistically significant positive relationship between RS trust and serendipity expectations across all platforms ($p < .05$) and for the aggregate model ($B = 0.240$, $p < .001$). These findings suggest that trust elevates expectations for serendipitous content. RS trust also had strong positive associations with satisfaction with serendipitous recommendations, both in each platform (all $p < .001$) and in the aggregate model ($B = 0.472$, $p < .001$). The effect strength was highest for LinkedIn and lowest for Spotify. H7 is strongly supported, highlighting trust as a key predictor of satisfaction.

**H8 and H9:** Another set of linear regressions revealed that curiosity significantly predicted serendipity expectations in the aggregate model ($B = 0.141$, $p = .032$) and for YouTube ($B = 0.211$, $p = .011$) but not other platforms. H8 is thus partially supported, suggesting curiosity has a modest role in shaping expectations. A final set of linear regressions was conducted to test whether curiosity predicts higher satisfaction with serendipitous recommendations. The results showed that curiosity significantly predicted satisfaction for YouTube and Netflix ($p < .05$) and in the aggregate model ($B = 0.135$, $p = .016$). No significant effects were observed on other platforms. These findings partially support H9, indicating that curiosity enhances satisfaction on visually immersive platforms.

# 5. Discussion

The analysis offered meaningful initial insights into the context of serendipitous content discovery on major digital platforms. This discussion section outlines the key findings, situates them within relevant literature, and addresses their implications and limitations.

## 5.1 Findings

The results of this study provide consistent support that serendipitous recommendation experiences fall short of user expectations across all examined platforms. The gap average for Amazon was the highest, followed by LinkedIn, even though their expectations averages were not the highest. These findings suggest a systematic under-delivery of surprise discovery, likely resulting from algorithmic biases toward familiarity and relevance optimization, as discussed in prior work by Pariser (2011, in Arguedas et al. (2022)), Reviglio (2019), and Krook (2025). Likewise, this confirms Kotkov et al.'s (2024) position that a gap persists between what users expect and experience from RSs.

Regression models showed that larger serendipity expectation-experience gaps were significantly associated with lower satisfaction, aligning well with the expectancy disconfirmation satisfaction paradigm. The effect was especially noticeable with Spotify and Netflix. Despite being rated as more 'discovery' overall and having smaller average gaps, these platforms elicited stronger negative effects when they failed to meet expectations. This could indicate higher baseline expectations for intelligent or personalized discovery on these platforms. Thus, the disappointment is amplified when expectations are not met, possibly due to the perceived failure of an otherwise "smart" system. Open-ended responses pointed out how users feel like they see duplicate content repeatedly and expressed frustration with being offered predictable or low-quality suggestions. This finding reinforces the importance of managing user expectations and balancing algorithmic personalization with discovery (Chen et al. 2019).

It was proposed that perceived control over the taste profile influences satisfaction, but this effect was only statistically significant for Spotify (marginally significant for the aggregate model). This could be attributed to users feeling a greater sense of agency on music platforms, where they actively engage with and shape their taste profiles through listening habits, playlists, and profiles, as Freeman, Gibbs, and Nansen (2022) suggest. The disappointment is more intensely felt when these actions do not translate into algorithmic adjustments. On other platforms, where personalization is more opaque, the lack of visible responsiveness to user signals may weaken the perception of control. The overall lack of satisfaction prediction from perceived control suggests that other factors, such as content quality or user traits, likely shape it.

Perceived serendipity experience and satisfaction varied significantly across platforms, supporting the view that platform design and content type shape serendipitous outcomes, as outlined by Smets et al. (2022), McCay-Peet, Toms, and Kelloway (2014; 2015), and Lutz, Hoffmann, and Meckel (2017). The findings also align with Nagulendra and Vassileva (2014), who emphasize that genuine discovery requires active user participation. While Netflix, Spotify, and YouTube are seen as passive “lean-back” environments, they appear to foster more serendipity, possibly through low-effort and immersive environments that fulfill the role of lightweight user participation. In contrast, the goal-driven design of LinkedIn and Amazon (Gorgoglione, Panniello, and Tuzhilin 2019) may limit exploratory behavior and, by extension, serendipitous discovery. The fact that users ranked these two platforms as the most repetitive ones of the study further supports this. This pattern also echoes Lutz, Hoffmann, and Meckel (2017), who found no positive association between serendipity and satisfaction in e-commerce contexts. Additionally, LinkedIn and Amazon were among the least-used platforms in the sample. As Reviglio (2019) notes, serendipity can emerge either through intentional system design or naturally through repeated user interaction. Infrequent usage could hinder both, reducing the system's ability to provide serendipitous content.

Trust in the RS significantly predicted both serendipity expectations and satisfaction across all platforms, replicating and extending previous research such as Pu, Chen, and Hu (2013). Across all platforms and in the aggregate model, the regression coefficients confirmed a dual role for trust – it raises users' expectations for discovery while enhancing their evaluation of serendipitous content. Notably, LinkedIn had the highest trust–satisfaction effect and the lowest overall trust scores, suggesting a polarizing experience: when trust is present, it matters significantly, but it may be harder to cultivate on this platform. Conversely, platforms like Spotify and YouTube, while still showing trust effects, may rely more on passive engagement or entertainment value, which might dilute the weight of trust relative to content quality or novelty. Still, some respondents in the open-ended item expressed confusion about how recommendation algorithms work and asked for more transparency.

Lastly, curiosity significantly predicted both serendipity expectations and satisfaction in the aggregate model, supporting its role as a general driver of openness to discovery. However, at the platform level, this effect was only significant for YouTube (both expectations and satisfaction) and Netflix (satisfaction only). This aligns with McCay-Peet, Toms, and Kelloway (2015) and Smets et al. (2022), who emphasize the role of content richness and exploratory affordances in activating curiosity. Platforms like YouTube and Netflix, which provide abundant, continuous content flows and interface elements that encourage browsing (e.g., autoplay, visual previews), appear to activate curiosity more effectively. In contrast, the other platforms may limit curiosity's impact by offering more structured or goal-directed experiences. These observations posit that satisfaction derived from curiosity only occurs when user traits align with platform affordances. They also build upon prior work (Chen et al. 2019; Stitini et al. 2023) that curiosity and broader interests can lead to increased satisfaction in contexts of novelty or uncertainty, while providing limited support for curiosity as a consistent predictor of serendipity expectations.

## 5.2 Implications

### 5.2.1 Practical Implications

The findings of this study point to a clear disconnect between user expectations and experiences of serendipity across popular digital platforms, which has significant implications for user satisfaction and long-term engagement with RSs. This section presents a practical guide and actionable steps for digital platforms to translate these findings into platform strategy and RS architecture, helping align personalization with user desires for surprise, exploration, and a sense of agency. The complete guide can be found in Appendix C. Practical Actions Guide.

Instead of viewing serendipity as a byproduct or fixed design component, platforms should consider it a dynamic capability that develops through the interaction between system design and user perception. The first phase of the guide emphasizes the foundational importance of internally defining serendipity and integrating it into platform strategy. As shown in the study, not all platforms are designed with the same goals; while content-rich and entertainment platforms can afford to push more serendipity, goal-oriented ones may require more subtle interventions.

Users consistently expected more serendipity than they experienced, and these gaps significantly reduced satisfaction, particularly on platforms where expectations for surprises were high (e.g., Spotify and Netflix). As a second phase to address these gaps and prevent long-term disengagement, platforms should assess expectations and experience, with this thesis serving as inspiration. Trust and curiosity emerged as key traits that shaped both expectations and satisfaction, reinforcing the importance of also collecting perceptual and behavioral data to identify serendipity-sensitive users.

The third phase focuses on improving user experience by building up trust, transparency, and control. Without a sense that the system understands and respects user preferences, attempts to introduce surprise may be interpreted as random or intrusive rather than intelligent or relevant. Similarly, the observed relationship between perceived control and satisfaction (though statistically significant only for Spotify and marginally for aggregate) suggests that agency over the taste profile plays a context-dependent but meaningful role in shaping RS evaluations. User-perceived agency should be accompanied by visible responsiveness in the RS to sustain engagement. Therefore, platforms should first focus on creating transparent and responsive recommendations where users can guide and understand personalization results. Ensuring the RS is perceived as trustworthy is more effective than simply implementing or increasing serendipity.

Serendipity is not equally valued or experienced but instead depends on both platform characteristics and individual traits. Platforms should, therefore, avoid static, one-size-fits-all approaches to serendipity and employ 'personalized diversification' (Malekzadeh Hamedani and Kaedi 2019) or adaptive serendipity. The fourth phase shifts to this technical implementation, with suggestions from literature-backed techniques (e.g., bandit algorithms or context-aware approaches) to labeled recommendation cues like "Hidden gem" or "Out of your usual pick." Additionally, offering user-adjustable discovery settings such as "safe," "balanced," and "surprise me" can provide flexibility for different curiosity levels and platform contexts. Interestingly, one respondent even suggested a similar exploratory feature: "a button that shows a version that pretends to know nothing of your preferences."

The final phase highlights the importance of continuous monitoring and iteration. As with any feature development, platforms would benefit from frequent direct user feedback to guide their RS serendipity design (Nalis et al. 2024; Stitini et al. 2023). Micro-feedback mechanisms (e.g., prompts like "Was this a good surprise?"), A/B testing of unexpected content delivery strategies, and routinely requesting feedback can help platforms recalibrate their RS to better match user desires without compromising relevance or usability. Over time, these behavioral signals and user feedback can inform the dynamic adaptation of serendipitous content delivery.

### 5.2.2 Theoretical Implications

This study contributes to the RS literature by separating and examining three interconnected constructs: expectations, experiences (perceptions), and satisfaction with serendipity. This distinction enables a more granular understanding of how users evaluate RSs beyond traditional metrics like accuracy or relevance.

The resulting gap between expected and perceived serendipity underscores the relevance of expectancy disconfirmation theory (EDT) in algorithmic environments. The study supports that unmet expectations, particularly regarding content serendipity, negatively affect satisfaction, suggesting that EDT provides a useful lens for understanding user experience with RSs, which has not been extensively acknowledged by prior research.

The differentiated roles of trust and curiosity offer further theoretical nuance. Trust can be inferred as an anticipatory *and* confirmatory variable, both shaping expectations and influencing satisfaction. This supports the idea that trust functions across the user journey and should be viewed as both a logical and emotional aspect in RS evaluation. Conversely, curiosity was associated more with satisfaction than expectations, suggesting its role may be more experiential and post hoc, supporting theories that link personality traits to openness toward discovery.

Lastly, platform context emerged as a crucial factor. The findings indicate that serendipity is not a fixed experience; instead, it is likely influenced by the RS or platform design, content type, and user interaction style. This extends literature emphasizing the need to understand the effects of RSs in specific contexts rather than treating serendipity as a static or universally desired outcome.

## 5.3 Limitations

Despite its contributions, this research faces several limitations. Due to self-selection and online recruitment, the sample had a younger demographic bias; this and uneven platform representation may limit the findings' generalizability. A reliance on self-reported data also introduces the potential biases of retrospective inaccuracies and central tendency effects on Likert scales. The study design limits the ability to draw causal conclusions, as it neither involved variable manipulation (e.g., creating control and variable groups with various platform changes) nor tracked longitudinal changes. Given that platform purposes and industries differ significantly, the cross comparisons of serendipity experience may partly reflect differences in user intent rather than RS quality alone. Additionally, constructs like "control" or "trust" can vary in interpretation depending on the platform context, affecting response consistency. Finally, confounding variables such as digital literacy, prior experience with RSs, or general attitudes toward personalization may have influenced the results.

### 5.4 Future Research

Future research could build on these findings by including behavioral data, such as browsing history or engagement metrics, to complement self-reported perceptions. Additionally, further research is required to assess the nature of expectations, as this study only covered the surface question of the extent to which they expect serendipity, but not why or from where they stem. In-depth research on each platform, platform industry (e.g., streaming, music, e-commerce, etc.), or demographic would address the limitation of generalization. Similarly, expanding the research to other industry RSs not covered (e.g., news and social media) could provide further valuable insights into how different contexts affect serendipity experience and satisfaction. Lastly, longitudinal studies could explore how user expectations and satisfaction with serendipity evolve over time and through repeated exposure to a platform's RS.

## 6. Conclusion

Recommender systems (RSs) have been around for many years, and after extensive academic focus on improving user satisfaction through non-accuracy measures such as serendipity, digital platforms are expected to deliver not just relevance but meaningful surprises. However, this thesis showed that while personalization has advanced, platforms often fall short of meeting user expectations for novel, diverse, and unexpected recommendations. The quantitative study examined how much users expect, perceive, and evaluate their serendipitous recommendations, and what other factors influence their satisfaction. A gap between expected and experienced serendipity was both common and significantly linked to lower satisfaction. The results further showed that satisfaction was shaped by factors such as curiosity, trust, and platform context. Visually rich platforms like YouTube and Netflix were more successful in converting curiosity into satisfaction, while goal-directed environments like Amazon and LinkedIn were seen as repetitive and less supportive of discovery. These findings contribute conceptual clarity to RS literature and offer a user-centered perspective on algorithmic serendipity.

For RS engineers, product managers, user experience professionals, and other practitioners, understanding user needs and expectations around serendipity is a crucial first step in the product or feature design process. The findings reinforce the importance of treating serendipity not as incidental, but as a strategic element of user experience. Aligning user expectations with system capabilities through transparency, feedback processes, exploratory design features, and visible responsiveness can enhance both trust and satisfaction. Furthermore, surprise discovery is not experienced equally across platforms or users, pointing to the need for more tailored interventions. Product teams should consider that what constitutes "serendipity" may vary by user, context, and platform purpose. In this light, serendipitous recommendation is not simply about surprise – it is about meaningful surprise that emerges from an environment users feel they can understand, influence, and trust.

## References


Afoudi, Yassine, Mohamed Lazaar, and Mohammed Al Achhab. 2021. “Hybrid Recommendation System Combined Content-Based Filtering and Collaborative Prediction Using Artificial Neural Network.” *Simulation Modelling Practice and Theory* 113 (December). https://doi.org/10.1016/j.simpat.2021.102375.

Amatriain, Xavier, and Justin Basilico. 2015. “Recommender Systems in Industry: A Netflix Case Study.” In *Recommender Systems Handbook*, edited by Francesco Ricci, Lior Rokach, and Bracha Shapira, 385–419. Springer. https://doi.org/10.1007/978-1-4899-7637-6_11.

Arguedas, Amy Ross, Craig T Robertson, Richard Fletcher, and Rasmus K Nielsen. 2022. “Echo Chambers, Filter Bubbles, and Polarisation: A Literature Review.” https://doi.org/10.60625/risj-etxj-7k60.

Armentano, Marcelo G., Roberto Abalde, Silvia Schiaffino, and Analia Amandi. 2014. “User Acceptance of Recommender Systems: Influence of the Preference Elicitation Algorithm.” In *Proceedings - 9th International Workshop on Semantic and Social Media Adaptation and Personalization, SMAP 2014*, 72–76. IEEE. https://doi.org/10.1109/SMAP.2014.18.

Armentano, Marcelo G., Ingrid Christensen, and Silvia Schiaffino. 2015. “Applying the Technology Acceptance Model to Evaluation of Recommender Systems.” *Polibits* 51 (January):73–79. https://doi.org/10.17562/pb-51-10.

Boudet, Julien, Lars Fiedler, Holger Hürtgen, and Kelsey Robinson. 2020. “Perspectives on Personalization @ Scale.” https://www.mckinsey.com/~/media/mckinsey/business%20functions/marketing%20and%20sales/our%20insights/perspectives%20on%20personalization%20at%20scale%20volume%202/perspectives-on-personalization-at-scale-volume-2-the-next-frontier.pdf.

Chen, Li, Marco De Gemmis, Alexander Felfernig, Pasquale Lops, Francesco Ricci, and Giovanni Semeraro. 2013. “Human Decision Making and Recommender Systems.” *ACM Transactions on Interactive Intelligent Systems* 3 (3): 17:1-17:7. https://doi.org/10.1145/2533670.2533675.

Chen, Li, Yonghua Yang, Ningxia Wang, Keping Yang, and Quan Yuan. 2019. “How Serendipity Improves User Satisfaction with Recommendations? A Large-Scale User Evaluation.” In *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, 240–50. Association for Computing Machinery, Inc. https://doi.org/10.1145/3308558.3313469.

Collins, Barry. 2020. “How To Stop Spotify Feeding You The Same Old Songs.” Forbes. June 11, 2020. https://www.forbes.com/sites/barrycollins/2020/06/11/how-to-stop-spotify-feeding-you-the-same-old-songs/.

Covington, Paul, Jay Adams, and Emre Sargin. 2016. “Deep Neural Networks for YouTube Recommendations.” In *RecSys ’16: Proceedings of the 10th ACM Conference on Recommender Systems*, 191–98. Association for Computing Machinery, Inc. https://doi.org/10.1145/2959100.2959190.

Davidson, James, Benjamin Liebald, Junning Liu, Palash Nandy, and Taylor Van Vleet. 2010. “The YouTube Video Recommendation System.” In *RecSys ’10: Proceedings of the 4th ACM Conference on Recommender Systems*, 293–96. Association for Computing Machinery, Inc. https://doi.org/10.1145/1864708.1864770.

Donaldson, Craig. 2024. “Amazon Using AI to Fix Its Broken Prime Video Algorithm.” Pocket-Lint. December 13, 2024. https://www.pocket-lint.com/prime-video-ai-topics/.

Fleder, Daniel, Kartik Hosanagar, and Andreas Buja. 2011. “Recommender Systems and Their Effects on Consumers.” In *MCIS 2011 Proceedings*, 59. AIS Electronic Library. http://aisel.aisnet.org/mcis2011/59.

Freeman, Sophie, Martin Gibbs, and Bjørn Nansen. 2022. “Don’t Mess with My Algorithm: Exploring the Relationship between Listeners and Automated Curation and Recommendation on Music Streaming Services.” *First Monday* 27 (1). https://doi.org/10.5210/fm.v27i1.11783.

Gomez-Uribe, Carlos A., and Neil Hunt. 2015. “The Netflix Recommender System: Algorithms, Business Value, and Innovation.” *ACM Transactions on Management Information Systems* 6 (4): 13:1-13:19. https://doi.org/10.1145/2843948.

Gorgoglione, Michele, Umberto Panniello, and Alexander Tuzhilin. 2019. “Recommendation Strategies in Personalization Applications.” *Information and Management* 56 (6). https://doi.org/10.1016/j.im.2019.01.005.

Hachey, Chris. 2025. “Retraining My YouTube Algorithm Saved My Sanity.” Pocket-Lint. February 21, 2025. https://www.pocket-lint.com/how-to-retrain-youtube/.

Harambam, Jaron, Mykola Makhortykh, Dimitrios Bountouridis, and Joris Van Hoboken. 2019. “Designing for the Better by Taking Users into Account: A Qualitative Evaluation of User Control Mechanisms in (News) Recommender Systems.” In *RecSys ’19: Proceedings of the 13th ACM Conference on Recommender Systems*, 69–77. Association for Computing Machinery, Inc. https://doi.org/10.1145/3298689.3347014.

Hardesty, Larry. 2019. “The History of Amazon’s Recommendation Algorithm.” Amazon Science. November 22, 2019. https://www.amazon.science/the-history-of-amazons-recommendation-algorithm?Utm_source=researcher_app&utm_medium=referral&utm_campaign=RESR_MRKT_Researcher_inbound.

Hazrati, Naieme, and Francesco Ricci. 2024. “Choice Models and Recommender Systems Effects on Users’ Choices.” *User Modeling and User-Adapted Interaction* 34 (1): 109–45. https://doi.org/10.1007/s11257-023-09366-x.

Janssens, Wim, Katrien Wijnen, Patrick De Pelsmacker, and Patrick Van Kenhove. 2008. *Marketing Research with SPSS*. Pearson Education. https://fdcfrearchworpress.wordpress.com/wp-content/uploads/2020/02/janssens-marketing-research-with-spss.pdf.

Javed, Aariz Faizan, and Syed Abdullah Ashraf. 2023. “Novelty In Recommender Systems for Effective Personalization in E-Commerce and Retail.” *Journal of Informatics Education and Research* 3 (2): 488–96. https://jier.org/index.php/journal/article/view/128.

Kim, Henry M., Bita Ghiasi, Max Spear, Marek Laskowski, and Jiye Li. 2017. “Online Serendipity: The Case for Curated Recommender Systems.” *Business Horizons* 60 (5): 613–20. https://doi.org/10.1016/j.bushor.2017.05.005.

Konstan, Joseph A., and John Riedl. 2012. “Recommender Systems: From Algorithms to User Experience.” *User Modeling and User-Adapted Interaction* 22 (March):101–23. https://doi.org/10.1007/s11257-011-9112-x.

Kotkov, Denis, Joseph A. Konstan, Qian Zhao, and Jari Veijalainen. 2018. “Investigating Serendipity in Recommender Systems Based on Real User Feedback.” In *Proceedings of the ACM Symposium on Applied Computing*, 1341–50. Association for Computing Machinery, Inc. https://doi.org/10.1145/3167132.3167276.

Kotkov, Denis, Alan Medlar, Triin Kask, and Dorota Glowacka. 2024. "The Dark Matter of Serendipity in Recommender Systems." In *CHIIR 2024 - Proceedings of the 2024 Conference on Human Information Interaction and Retrieval*, 108–18. Association for Computing Machinery, Inc. https://doi.org/10.1145/3627508.3638342.

Krook, Joshua. 2025. "Building Serendipity into Recommender Algorithms on Online Platforms: Reviving the Chaos and Randomness of the Early Internet Aesthetic." *Law, Technology and Humans*, January. https://doi.org/10.5204/lthj.3730.

Li, Qinglong, Il-Young Choi, and Jae-Kyeong Kim. 2020. "Evaluation of Recommendation System for Sustainable E-Commerce: Accuracy, Diversity and Customer Satisfaction." *Preprints*, January. https://doi.org/10.20944/preprints202001.0015.v1.

Linden, Greg, Brent Smith, and Jeremy York. 2003. "Amazon.Com Recommendations: Item-to-Item Collaborative Filtering." *IEEE Internet Computing* 7 (1): 76–80. https://doi.org/10.1109/MIC.2003.1167344.

Lutz, Christoph, Christian Pieter Hoffmann, and Miriam Meckel. 2017. "Online Serendipity: A Contextual Differentiation of Antecedents and Outcomes." *Journal of the Association for Information Science and Technology* 68 (7): 1698–1710. https://doi.org/10.1002/asi.23771.

Lyngs, Ulrik, Reuben Binns, Max Van Kleek, and Nigel Shadbolt. 2018. "'So, Tell Me What Users Want, What They Really, Really Want!'" In *Conference on Human Factors in Computing Systems - Proceedings*. Vol. 2018-April. Association for Computing Machinery, Inc. https://doi.org/10.1145/3170427.3188397.

Malekzadeh Hamedani, Elaheh, and Marjan Kaedi. 2019. "Recommending the Long Tail Items through Personalized Diversification." *Knowledge-Based Systems* 164 (January):348–57. https://doi.org/10.1016/j.knosys.2018.11.004.

Martínez-López, Francisco J., Inma Rodríguez-Ardura, Juan Carlos Gázquez-Abad, Manuel J. Sánchez-Franco, and Claudia C. Cabal. 2010. "Psychological Elements Explaining the Consumer's Adoption and Use of a Website Recommendation System: A Theoretical Framework Proposal." *Internet Research* 20 (3): 316–41. https://doi.org/10.1108/10662241011050731.

McCay-Peet, Lori, Elaine G. Toms, and E. Kevin Kelloway. 2014. "Development and Assessment of the Content Validity of a Scale to Measure How Well a Digital Environment Facilitates Serendipity." *Information Research - An International Electronic Journal* 19 (3). https://www.informationr.net/ir/19-3/paper630.html#.VNpgJlpjonM.

———. 2015. "Examination of Relationships among Serendipity, the Environment, and Individual Differences." *Information Processing and Management* 51 (4): 391–412. https://doi.org/10.1016/j.ipm.2015.02.004.

McInerney, James, Benjamin Lacker, Samantha Hansen, Karl Higley, Hugues Bouchard, Alois Gruson, and Rishabh Mehrotra. 2018. "Explore, Exploit, and Explain: Personalizing Explainable Recommendations with Bandits." In *RecSys '18: Proceedings of the 12th ACM Conference on Recommender Systems*, 31–39. Association for Computing Machinery, Inc. https://doi.org/10.1145/3240323.3240354.

Merriam-Webster.com Dictionary s.v. n.d. "Serendipity." Accessed April 29, 2025. https://www.merriam-webster.com/dictionary/serendipity.

Mudaliyar, Mylie. 2024. "Behind the Playlist: Analyzing Spotify's Recommendation System." https://medium.com/@myliemudaliyar/behind-the-playlist-analyzing-spotifys-recommendation-system-5044a13f5ccf.

Nagulendra, Sayooran, and Julita Vassileva. 2014. "Understanding and Controlling the Filter Bubble through Interactive Visualization: A User Study." In *HT 2014 - Proceedings of the 25th ACM Conference on Hypertext and Social Media*, 107–15. Association for Computing Machinery, Inc. https://doi.org/10.1145/2631775.2631811.

Nalis, Irina, Tobias Sippl, Thomas Elmar Kolb, and Julia Neidhardt. 2024. "Navigating Serendipity - An Experimental User Study on the Interplay of Trust and Serendipity in Recommender Systems." In *UMAP 2024 - Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*, 386–93. Association for Computing Machinery, Inc. https://doi.org/10.1145/3631700.3664901.

Nenty, H. Johnson. 2009. "Writing a Quantitative Research Thesis." *International Journal of Educational Sciences* 1 (1): 19–32. https://doi.org/10.1080/09751122.2009.11889972.

Olsson, Thomas. 2014. "Layers of User Expectations of Future Technologies: An Early Framework." In *Conference on Human Factors in Computing Systems - Proceedings*, 1957–62. Association for Computing Machinery, Inc. https://doi.org/10.1145/2559206.2581225.

Ping, Yanni, Yang Li, and Jiaxin Zhu. 2024. "Beyond Accuracy Measures: The Effect of Diversity, Novelty and Serendipity in Recommender Systems on User Engagement." *Electronic Commerce Research*, February. https://doi.org/10.1007/s10660-024-09813-w.

Pu, Pearl, Li Chen, and Rong Hu. 2013. "A User-Centric Evaluation Framework for Recommender Systems." In *RecSys '11: Proceedings of the Fifth ACM Conference on Recommender Systems*, 157–64. Association for Computing Machinery, Inc. https://doi.org/10.1145/2043932.2043962.

Reviglio, Urbano. 2019. "Serendipity as an Emerging Design Principle of the Infosphere: Challenges and Opportunities." *Ethics and Information Technology* 21 (2): 151–66. https://doi.org/10.1007/s10676-018-9496-y.

Ricci, Francesco, Lior Rokach, and Bracha Shapira. 2011. "Recommender Systems Handbook: Introduction and Challenges." In *Recommender Systems Handbook*, 1–35. Springer US. https://doi.org/10.1007/978-0-387-85820-3_1.

Smets, Annelien, Lien Michiels, Toine Bogers, and Lennart Björneborn. 2022. "Serendipity in Recommender Systems Beyond the Algorithm: A Feature Repository and Experimental Design." In *9th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems, IntRS 2022*, 3222:46–66. CEUR Workshop Proceedings. https://vbn.aau.dk/en/publications/serendipity-in-recommender-systems-beyond-the-algorithm-a-feature.

Stitini, Oumaima, Iván García-Magariño, Soulaimane Kaloun, and Omar Bencharef. 2023. "Towards Ideal and Efficient Recommendation Systems Based on the Five Evaluation Concepts Promoting Serendipity." *Journal of Advances in Information Technology* 14 (4): 701–17. https://doi.org/10.12720/jait.14.4.701-717.

"The Jamovi Project." 2022. jamovi. (Version 2.3) [Computer Software]. https://www.jamovi.org.

Wang, Yen-Yao, Anthony Townsend, Andy Luse, and Brian Mennecke. 2012. "The Determinants of Acceptance of Recommender Systems: Applying the UTAUT Model." In *AMCIS 2012 Proceedings*. AIS Electronic Library. http://aisel.aisnet.org/amcis2012/proceedings/EBusiness/2.

Wanjala, Alvin. 2024. “Here’s How I Reset My Netflix Recommendations When They Aren’t Good Anymore.” Make Use Of. September 26, 2024. https://www.makeuseof.com/tag/how-to-recalibrate-your-netflix-recommendations/.

Wasco, Gabrielle, and Min Read. 2024. “Chained to the AlgoRhythm: How Spotify Standardizes Our Listening.” *The Decision Lab*. https://thedecisionlab.com/insights/consumer-insights/chained-to-the-algorhythm-how-spotify-standardizes-our-listening.

Wu, Lili, Sam Shah, Sean Choi, Mitul Tiwari, and Christian Posse. 2014. “The Browsemaps: Collaborative Filtering at LinkedIn.” *CEUR Workshop Proceedings* 1271 (January). https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=08ce55006b64b57a367bcedd768785abeff0c191.

Zanker, Markus, Francesco Ricci, Dietmar Jannach, and Loren Terveen. 2010. “Measuring the Impact of Personalization and Recommendation on User Behaviour.” *International Journal of Human Computer Studies* 68 (8): 469–71. https://doi.org/10.1016/j.ijhcs.2010.04.002.

Ziarani, Reza Jafari, and Reza Ravanmehr. 2021a. “Deep Neural Network Approach for a Serendipity-Oriented Recommendation System.” *Expert Systems with Applications* 185 (December). https://doi.org/10.1016/j.eswa.2021.115660.

———. 2021b. “Serendipity in Recommender Systems: A Systematic Literature Review.” *Journal of Computer Science and Technology* 36 (2): 375–96. https://doi.org/10.1007/s11390-020-0135-9.

# Appendix

## A. Abbreviations

| | |
|---|---|
| AI | Artificial intelligence |
| CBF | Content-based filtering |
| CEI-II | Curiosity and exploration inventory-II |
| CF | Collaborative filtering |
| EDT | Expectancy disconfirmation theory |
| ResQue | Recommender system's quality of user experience |
| RS | Recommender system |
| RSs | Recommender systems |
| TAM | Technology acceptance model |
| UTAUT | Unified theory of acceptance and use of technology |

## B. Tables

*Appendix B Table 2. Survey Items*

| Construct | Item | Measurement (per chosen platform) | Source adaptation |
|---|---|---|---|
| **Engagement context** | | | |
| Platform choice | Which of the following digital platforms do you use regularly (at least once per month)? | Multi-choice: Netflix, YouTube, Spotify, Amazon, LinkedIn | Author-original |
| Usage frequency | How often do you use the platform? | Single choice: Once a month, A few times a month, Once a week, 2-3 times a week, 4-6 times a week, Daily | Author-original |
| Search behavior | Usually (at least 50% of the time) what are you searching for? | Single choice: Open search, Vague search, Specific item | Author-original |
| Awareness | Are you aware that the platform tracks your item history and usage, and uses that data to construct a 'taste profile' and make personalized recommendations? | 3-point Likert: Not aware, Somewhat aware, Aware | Author-original |
| **Expectations vs. experience**** | | | |
| Serendipity 1 | Through recommendations, I make an *accidental* fortunate discovery of content that is useful for me (no search intention, recommendation is unplanned or unprompted). | 5-point agreement Likert | Lutz, Hoffmann, and Meckel (2017) |
| Serendipity 2 | Through recommendations, I make an *unexpected* fortunate discovery of content that is useful for me (deviates from expectations within a specific search). | | Lutz, Hoffmann, and Meckel (2017) |
| Serendipity 3 | I encounter useful information, ideas, or resources (recommendations) that I am not looking for. | | Lutz, Hoffmann, and Meckel (2017) |
| Control | The platform allows me to modify and control my taste profile (the collection of data from past item usage to make personalized recommendations). | | Pu, Chen, and Hu (2013) |
| **Satisfaction** | | | |
| Satisfaction 1 | Generally, I am satisfied with the novelty (newness, never seen before) of recommendations. | 5-point satisfaction Likert | Pu, Chen, and Hu (2013) |
| Satisfaction 2 | Generally, I am satisfied with the diversity (variation) of recommendations. | | Pu, Chen, and Hu (2013) |
| Satisfaction 3 | Generally, I am satisfied with how often I make an accidental or unexpected fortunate discovery through recommendations. | | Pu, Chen, and Hu (2013) |
| **Trust** | | | |
| Trust 1 | The platform recommendation system can be trusted. | 5-point agreement Likert | Pu, Chen, and Hu (2013) |
| Trust 2 | The platform recommendation system makes me more confident about my item selection/decision. | | Pu, Chen, and Hu (2013) |
| **Curiosity ("Both online and offline, I…")** | | | |
| Curiosity 1 | ...am on the lookout for new things and experiences. | 5-point agreement Likert | Chen et al. (2019) |
| Curiosity 2 | ...am always looking for experiences that challenge how I think about myself and the world. | | Chen et al. (2019) |

| | | | |
|---|---|---|---|
| Curiosity 3 | ...prefer situations that are excitingly unpredictable. | | Chen et al. (2019) |
| **Comments and demographics** | | | |
| Balance | Please rate how you find the balance between repetitive/similar and new/discovery recommendations. | 5-point Likert: More repetitive, Somewhat repetitive, Balanced, Somewhat discovery, More discovery | Author-original |
| Comment | Is there anything else you'd like to share about your experience with recommendations on the platform(s)? Any feedback to the platform(s) on how they could do better? | Open answer box | Author-original |
| Nationality | What is your nationality? | Open answer box | Pu, Chen, and Hu (2013) |
| Age | How old are you? | Single choice: Under 18, 18-24, 25-34, 35-44, 45-54, 55+ | Pu, Chen, and Hu (2013) |
| Gender | What gender do you identify as? | Single choice: Male, Female, Non-binary, Prefer not to say | Pu, Chen, and Hu (2013) |

*Note*. **Each statement is followed by two answering sections side-by-side, one "I expect this from the platform.", and two "I often experience this on the platform."

*Appendix B Table 3. Cronbach's Alpha Reliability Analysis*

| **Construct** | **Platform** | **Mean** | **Standard Deviation** | **Cronbach's a** |
|---|---|---|---|---|
| Serendipity Expectation (n=3) | Netflix | 3.81 | 0.841 | 0.614 |
| | YouTube | 4.02 | 0.827 | 0.687 |
| | Spotify | 3.81 | 0.904 | 0.763 |
| | Amazon | 3.60 | 0.915 | 0.734 |
| | LinkedIn | 3.57 | 0.967 | 0.726 |
| Serendipity Experience (n=3) | Netflix | 3.54 | 0.889 | 0.712 |
| | YouTube | 3.92 | 0.856 | 0.725 |
| | Spotify | 3.58 | 0.906 | 0.709 |
| | Amazon | 3.33 | 0.962 | 0.720 |
| | LinkedIn | 3.30 | 0.941 | 0.734 |
| Satisfaction (n=3) | Netflix | 3.53 | 0.901 | 0.838 |
| | YouTube | 3.85 | 0.849 | 0.809 |
| | Spotify | 3.75 | 0.825 | 0.825 |
| | Amazon | 3.41 | 0.797 | 0.753 |
| | LinkedIn | 3.24 | 0.922 | 0.838 |
| Trust (n=2) | Netflix | 3.55 | 0.902 | 0.719 |
| | YouTube | 3.60 | 0.878 | 0.725 |
| | Spotify | 3.67 | 0.862 | 0.747 |
| | Amazon | 3.34 | 0.846 | 0.655 |
| | LinkedIn | 3.45 | 0.899 | 0.716 |
| Curiosity (n=3) | | 3.65 | 0.754 | 0.657 |

*Appendix B Table 4. Sample Descriptive Statistics*

| Construct | Choice | N | Percent |
|---|---|---|---|
| Platform | Netflix | 135 | 64.9% |
| | YouTube | 179 | 86.1% |
| | Spotify | 140 | 67.3% |
| | Amazon | 88 | 42.3% |
| | LinkedIn | 115 | 55.3% |
| Age | Under 18 | 2 | 1.0% |
| | 18-24 | 121 | 58.2% |
| | 25-34 | 56 | 26.9% |
| | 35-44 | 22 | 10.6% |
| | 45-54 | 5 | 2.4% |
| | 55+ | 2 | 1.0% |
| Gender | Male | 69 | 33.2% |
| | Female | 132 | 63.5% |
| | Non-binary | 4 | 1.9% |
| | Prefer not to say | 3 | 1.4% |
| Nationality | Europe | 114 | 54.8% |
| | North America | 41 | 19.7% |
| | Asia | 32 | 15.4% |
| | Africa | 10 | 4.8% |
| | Oceania | 4 | 1.9% |
| | South America | 3 | 1.4% |
| | N/A | 4 | 1.9% |

*Appendix B Table 5. Subsample Descriptive Statistics*

| | | ***Netflix (n=135)*** | | ***YouTube (n=179)*** | | ***Spotify (n=140)*** | | ***Amazon (n=88)*** | | ***LinkedIn (n=115)*** | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Construct** | **Choice** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** |
| Usage frequency | Once a month | 4 | 3.0% | 4 | 2.2% | 2 | 1.4% | 16 | 18.2% | 7 | 6.1% |
| | A few times a month | 26 | 19.3% | 13 | 7.3% | 5 | 3.6% | 30 | 34.1% | 18 | 15.7% |
| | Once a week | 15 | 11.1% | 7 | 3.9% | 4 | 2.9% | 17 | 19.3% | 23 | 20.0% |
| | 2-3 times a week | 38 | 28.1% | 30 | 16.8% | 15 | 10.7% | 14 | 15.9% | 18 | 15.7% |
| | 4-6 times a week | 22 | 16.3% | 29 | 16.2% | 22 | 15.7% | 8 | 9.1% | 10 | 8.7% |
| | Daily | 30 | 22.2% | 96 | 53.6% | 92 | 65.7% | 3 | 3.4% | 39 | 33.9% |
| Search behavior | Open search | 45 | 33.3% | 68 | 38.0% | 34 | 24.3% | 13 | 14.8% | 48 | 41.7% |
| | Vague search | 45 | 33.3% | 46 | 25.7% | 40 | 28.6% | 15 | 17.0% | 30 | 26.1% |
| | Specific item | 45 | 33.3% | 65 | 36.3% | 66 | 47.1% | 60 | 68.2% | 37 | 32.2% |
| Awareness | Not aware | 5 | 3.7% | 4 | 2.2% | 3 | 2.1% | 2 | 2.3% | 9 | 7.8% |
| | Somewhat aware | 17 | 12.6% | 15 | 8.4% | 10 | 7.1% | 12 | 13.6% | 33 | 28.7% |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Aware | 113 | 83.7% | 160 | 89.4% | 127 | 90.7% | 74 | 84.1% | 73 | 63.5% |
| Balance | More repetitive | 12 | 8.9% | 10 | 5.6% | 11 | 7.9% | 9 | 10.2% | 15 | 13.0% |
| | Somewhat repetitive | 36 | 26.7% | 46 | 25.7% | 28 | 20.0% | 21 | 23.9% | 28 | 24.3% |
| | Balanced | 37 | 27.4% | 51 | 28.5% | 38 | 27.1% | 28 | 31.8% | 29 | 25.2% |
| | Somewhat discovery | 40 | 29.6% | 48 | 26.8% | 44 | 31.4% | 25 | 28.4% | 32 | 27.8% |
| | More discovery | 10 | 7.4% | 24 | 13.4% | 19 | 13.6% | 5 | 5.7% | 11 | 9.6% |
| Age | Under 18 | 1 | 0.7% | 1 | 0.6% | 1 | 0.7% | 0 | 0.0% | 0 | 0.0% |
| | 18-24 | 74 | 54.8% | 104 | 58.1% | 92 | 65.7% | 47 | 53.4% | 73 | 63.5% |
| | 25-34 | 39 | 28.9% | 48 | 26.8% | 35 | 25.0% | 19 | 21.6% | 26 | 22.6% |
| | 35-44 | 15 | 11.1% | 20 | 11.2% | 11 | 7.9% | 17 | 19.3% | 12 | 10.4% |
| | 45-54 | 4 | 3.0% | 4 | 2.2% | 1 | 0.7% | 3 | 3.4% | 3 | 2.6% |
| | 55+ | 2 | 1.5% | 2 | 1.1% | 0 | 0.0% | 2 | 2.3% | 1 | 0.9% |
| Gender | Male | 42 | 31.1% | 61 | 34.1% | 46 | 32.9% | 25 | 28.4% | 38 | 33.0% |
| | Female | 88 | 65.2% | 111 | 62.0% | 90 | 64.3% | 60 | 68.2% | 76 | 66.1% |
| | Non-binary | 4 | 3.0% | 4 | 2.2% | 3 | 2.1% | 2 | 2.3% | 0 | 0.0% |
| | Prefer not to say | 1 | 0.7% | 3 | 1.7% | 1 | 0.7% | 1 | 1.1% | 1 | 0.9% |

*Appendix B Table 6. Summary of Open-Ended Comments*

| Theme | Description | Quotes | Key Insight |
|---|---|---|---|
| 1. Lack of serendipity and diversity | Frustration with repetitive or overly similar recommendations that limit exploration of new content. | • Spotify often recommends the same kind of music... nothing really new is being injected<br>• YouTube just repeats the same videos on my recommended even after refreshing the page<br>• All these platforms are predictive of past behaviors, making it hard for new and exciting items to be suggested<br>• Should be able to alter interests as years go by<br>• LinkedIn doesn't recommend enough new resources or experiences<br>• Spotify didn't capture my musical taste and it's been more than 4 years and I listen on it daily | Users want novelty but feel algorithms overemphasize past behavior at the expense of genuine exploration, feeling even with time it doesn't get better. |
| 2. Algorithm transparency and control | Confusion or lack of trust in how algorithms work and a desire for more user control over recommendations. | • I do not understand how the algorithm works and wish I could use it more to my advantage (especially for Spotify)<br>• On YouTube, I can see how popular the content is with other people which gives me confidence in trying the new thing<br>• On Netflix, there is no direct way to verify the general "approval" for the new content and I often check ratings of content on different sites<br>• For certain platforms, I have that recommended features for tracking turned off | There's a demand for explainability and user agency, otherwise they distrust and avoid the system. |

| | | | |
|---|---|---|---|
| 3. Content quality and relevance | Critiques of low-quality, irrelevant, or unsafe content being recommended, particularly commercial-related. | • The recommendations [on Amazon] that I get are obviously for similar items, however, they often deviate from my original purchase in terms of quality<br>• Seems the algorithm pushes Netflix's agenda<br>• YouTube recommendations fill up with irrelevant content<br>• YouTube just has so much misinformation that it's basically never safe to click on a recommended video<br>• The recommendations are not optimizing the user experience for the user's betterment, but to capture for sales or ads | Users want trustworthy, relevant content, not just high-engagement or monetized suggestions. |
| 4. Openness and reflection | Positive perceptions across platforms, with some openness to both discovery and personalization. | • Netflix has done a better job of balance recently<br>• I love journeys on YouTube, being led from one thing to another ... Amazon can't compete with this<br>• Perhaps if I interacted a little more with LinkedIn, I would allow it to "get to know me" a little better | Users are noticing changes over time and emitting positive perceptions and self-reflection. |
| 5. Missing features and suggested improvements | Users propose specific features to enhance both personalization and discovery. | • Make it clearer what is a recommendation and what is an actual search from the user<br>• A button that shows a version that pretends to know nothing of your preferences could be an interesting feature<br>• Include a quality filter for suggestions (esp. Amazon)<br>• YouTube should balance the amount of new and old contents recommendation across all types of content | Suggested features reflect users' need for both better personalization and mechanisms to encourage discovery, such as toggling filters or resetting preferences. |

*Appendix B Table 7. Wilcoxon Signed-Rank Results for H1: Serendipity Expectations vs. Experience*

| | **Expectations** | | **Experience** | | | |
|---|---|---|---|---|---|---|
| **Platform** | **Mean** | **SE** | **Mean** | **SE** | **Wilcoxon W** | **p** |
| Netflix | 3.81 | 0.0724 | 3.54 | 0.0765 | 3011 | < .001 |
| YouTube | 3.92 | 0.0618 | 3.67 | 0.0649 | 4760 | 0.047 |
| Spotify | 3.81 | 0.0764 | 3.58 | 0.0766 | 3323 | < .001 |
| Amazon | 3.60 | 0.1025 | 3.33 | 0.1076 | 1233 | 0.002 |
| LinkedIn | 3.57 | 0.0878 | 3.30 | 0.0877 | 2727 | 0.002 |
| All Platforms | 3.74 | 0.0494 | 3.56 | 0.0466 | 10483 | < .001 |

*Appendix B Table 8. Regression Results for H2: Serendipity Gap Predicting Satisfaction*

| **Platform** | | **B** | **SE** | **t** | **p** | **R²** |
|---|---|---|---|---|---|---|
| Netflix | (Constant) | 3.648 | 0.077 | 47.43 | | |
| | Serendipity Gap | –0.438 | 0.097 | –4.53 | < .001 | 0.134 |
| YouTube | (Constant) | 3.869 | 0.063 | 61.20 | | |
| | Serendipity Gap | –0.189 | 0.084 | –2.25 | 0.025 | 0.028 |
| Spotify | (Constant) | 3.865 | 0.063 | 61.29 | | |
| | Serendipity Gap | –0.476 | 0.071 | –6.66 | < .001 | 0.243 |
| Amazon | (Constant) | 3.47 | 0.089 | 39.38 | | |
| | Serendipity Gap | –0.226 | 0.107 | –2.12 | 0.037 | 0.050 |
| LinkedIn | (Constant) | 3.298 | 0.089 | 37.21 | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Serendipity Gap | –0.217 | 0.099 | –2.19 | 0.030 | 0.041 |
| All Platforms | (Constant) | 3.635 | 0.043 | 85.36 | | |
| | Serendipity Gap | –0.233 | 0.065 | –3.58 | < .001 | 0.059 |

*Appendix B Table 9. Regression Results for H3: Control Gap Predicting Satisfaction*

| Platform | | B | SE | t | p | R² |
|---|---|---|---|---|---|---|
| Netflix | (Constant) | 3.575 | 0.082 | 43.55 | | |
| | Serendipity Gap | –0.102 | 0.066 | –1.56 | 0.121 | 0.018 |
| YouTube | (Constant) | 3.866 | 0.067 | 57.43 | | |
| | Serendipity Gap | –0.040 | 0.059 | –0.69 | 0.489 | 0.003 |
| Spotify | (Constant) | 3.864 | 0.076 | 50.68 | | |
| | Serendipity Gap | –0.196 | 0.063 | –3.11 | 0.002 | 0.066 |
| Amazon | (Constant) | 3.445 | 0.089 | 38.57 | | |
| | Serendipity Gap | –0.087 | 0.069 | –1.25 | 0.214 | 0.018 |
| LinkedIn | (Constant) | 3.302 | 0.091 | 36.16 | | |
| | Serendipity Gap | –0.154 | 0.083 | –1.86 | 0.065 | 0.030 |
| All Platforms | (Constant) | 3.625 | 0.045 | 80.64 | | |
| | Serendipity Gap | –0.084 | 0.043 | –1.97 | 0.051 | 0.018 |

*Appendix B Table 10. Friedman Test for H4: Serendipity Experience Difference by Platform*

| Test | χ² | df | p |
|---|---|---|---|
| Friedman | 15.6 | 4 | 0.004 |

| Pairwise Comparisons (Durbin-Conover) | | Statistic | p |
|---|---|---|---|
| Pair 1 | Netflix - YouTube | 0.0886 | 0.930 |
| Pair 2 | Netflix - Spotify | 0.9750 | 0.331 |
| Pair 3 | Netflix - Amazon | 2.1273 | 0.035 |
| Pair 4 | Netflix - LinkedIn | 3.4569 | <.001 |
| Pair 5 | YouTube - Spotify | 0.8864 | 0.377 |
| Pair 6 | YouTube - Amazon | 2.0387 | 0.044 |
| Pair 7 | YouTube - LinkedIn | 3.3682 | <.001 |
| Pair 8 | Spotify - Amazon | 1.1523 | 0.251 |
| Pair 9 | Spotify - LinkedIn | 2.4819 | 0.014 |
| Pair 10 | Amazon - LinkedIn | 1.3296 | 0.186 |

*Appendix B Table 11. Friedman Test for H5: Serendipity Satisfaction Difference by Platform*

| Test | χ² | df | p |
|---|---|---|---|
| Friedman | 16.5 | 4 | 0.002 |

| Pairwise Comparisons (Durbin-Conover) | | Statistic | p |
|---|---|---|---|
| Pair 1 | Netflix - YouTube | 1.158 | 0.249 |
| Pair 2 | Netflix - Spotify | 1.713 | 0.089 |
| Pair 3 | Netflix - Amazon | 1.343 | 0.182 |

| | | | |
|---|---|---|---|
| Pair 4 | Netflix - LinkedIn | 1.760 | 0.081 |
| Pair 5 | YouTube - Spotify | 0.556 | 0.579 |
| Pair 6 | YouTube - Amazon | 2.500 | 0.014 |
| Pair 7 | YouTube - LinkedIn | 2.917 | 0.004 |
| Pair 8 | Spotify - Amazon | 3.056 | 0.003 |
| Pair 9 | Spotify - LinkedIn | 3.473 | <.001 |
| Pair 10 | Amazon - LinkedIn | 0.417 | 0.678 |

*Appendix B Table 12. Regression Results for H6: Trust Predicting Serendipity Expectations*

| **Platform** | | **B** | **SE** | **t** | **p** | **R²** |
|---|---|---|---|---|---|---|
| Netflix | (Constant) | 2.779 | 0.281 | 9.88 | | |
| | Trust | 0.290 | 0.077 | 3.77 | < .001 | 0.097 |
| YouTube | (Constant) | 2.839 | 0.246 | 11.55 | | |
| | Trust | 0.327 | 0.066 | 4.93 | < .001 | 0.121 |
| Spotify | (Constant) | 2.752 | 0.324 | 8.51 | | |
| | Trust | 0.289 | 0.086 | 3.37 | < .001 | 0.076 |
| Amazon | (Constant) | 2.399 | 0.379 | 6.33 | | |
| | Trust | 0.360 | 0.110 | 3.28 | 0.002 | 0.111 |
| LinkedIn | (Constant) | 2.073 | 0.330 | 6.28 | | |
| | Trust | 0.433 | 0.093 | 4.67 | < .001 | 0.162 |
| All Platforms | (Constant) | 2.895 | 0.242 | 11.96 | | |
| | Trust | 0.240 | 0.068 | 3.55 | < .001 | 0.058 |

*Appendix B Table 13. Regression Results for H7: Trust Predicting Serendipity Satisfaction*

| **Platform** | | **B** | **SE** | **t** | **p** | **R²** |
|---|---|---|---|---|---|---|
| Netflix | (Constant) | 1.255 | 0.243 | 5.16 | | |
| | Trust | 0.641 | 0.067 | 9.66 | < .001 | 0.412 |
| YouTube | (Constant) | 1.782 | 0.217 | 8.23 | | |
| | Trust | 0.575 | 0.058 | 9.83 | < .001 | 0.353 |
| Spotify | (Constant) | 1.835 | 0.257 | 7.14 | | |
| | Trust | 0.523 | 0.068 | 7.67 | < .001 | 0.299 |
| Amazon | (Constant) | 1.474 | 0.276 | 5.34 | | |
| | Trust | 0.579 | 0.080 | 7.23 | < .001 | 0.378 |
| LinkedIn | (Constant) | 0.774 | 0.246 | 3.14 | | |
| | Trust | 0.715 | 0.069 | 10.34 | < .001 | 0.486 |
| All Platforms | (Constant) | 1.940 | 0.177 | 10.96 | | |
| | Trust | 0.472 | 0.050 | 9.52 | < .001 | 0.306 |

*Appendix B Table 14. Regression Results for H8: Curiosity Predicting Serendipity Expectations*

| **Platform** | | **B** | **SE** | **t** | **p** | **R²** |
|---|---|---|---|---|---|---|
| Netflix | (Constant) | 3.306 | 0.408 | 8.10 | | |
| | Curiosity | 0.131 | 0.105 | 1.25 | 0.214 | 0.0116 |
| YouTube | (Constant) | 3.250 | 0.303 | 10.72 | | |
| | Curiosity | 0.211 | 0.082 | 2.58 | 0.011 | 0.0363 |
| Spotify | (Constant) | 3.213 | 0.398 | 8.08 | | |
| | Curiosity | 0.163 | 0.106 | 1.53 | 0.127 | 0.0168 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Amazon | (Constant) | 2.688 | 0.521 | 5.16 | | |
| | Curiosity | 0.247 | 0.138 | 1.78 | 0.078 | 0.0357 |
| LinkedIn | (Constant) | 3.552 | 0.517 | 6.86 | | |
| | Curiosity | 0.004 | 0.138 | 0.03 | 0.979 | <0.001 |
| All Platforms | (Constant) | 3.222 | 0.242 | 13.29 | | |
| | Curiosity | 0.141 | 0.065 | 2.16 | 0.032 | 0.0222 |

*Appendix B Table 15. Regression Results for H9: Curiosity Predicting Serendipity Satisfaction*

| **Platform** | | **B** | **SE** | **t** | **p** | **R²** |
|---|---|---|---|---|---|---|
| Netflix | (Constant) | 2.536 | 0.431 | 5.89 | | |
| | Curiosity | 0.261 | 0.111 | 2.35 | 0.020 | 0.0397 |
| YouTube | (Constant) | 2.860 | 0.308 | 9.29 | | |
| | Curiosity | 0.273 | 0.083 | 3.29 | 0.001 | 0.0575 |
| Spotify | (Constant) | 3.667 | 0.366 | 10.02 | | |
| | Curiosity | 0.024 | 0.098 | 0.25 | 0.806 | < .001 |
| Amazon | (Constant) | 3.065 | 0.461 | 6.66 | | |
| | Curiosity | 0.093 | 0.122 | 0.76 | 0.449 | 0.0067 |
| LinkedIn | (Constant) | 2.792 | 0.491 | 5.68 | | |
| | Curiosity | 0.121 | 0.131 | 0.93 | 0.356 | <0.001 |
| All Platforms | (Constant) | 3.101 | 0.206 | 15.05 | | |
| | Curiosity | 0.135 | 0.056 | 2.44 | 0.016 | 0.0281 |

# C. Practical Actions Guide

<table>
<tr><th colspan="2">Phase 1. Diagnose Platform Readiness</th></tr>
<tr><td>Define serendipity internally</td><td><ul><li>Conduct workshops (product, user experience, data science, marketing) to co-define what “serendipitous discovery” looks like in your context</li><li>Integrate it into platform strategy<ul><li>Exploratory platform: serendipity strategy can be more aggressive</li><li>Utilitarian platform: strategy requires more subtle, context-aware interventions</li></ul></li></ul></td></tr>
<tr><td>Run a technical audit</td><td><ul><li>Evaluate current (or future) RS architecture: type, data availability, and algorithm training capabilities</li><li>Identify signs of overspecialization (e.g., repetition rate, diversity metrics)</li><li>Check visibility of algorithm logic, control options, and current user feedback loops</li><li>Check what key metrics are being measured and how: diversity, novelty, trust, etc.</li></ul></td></tr>
<tr><th colspan="2">Phase 2. Understand Your Users</th></tr>
<tr><td>Measure expectation-experience gaps</td><td><ul><li>Deploy short-form, Likert surveys with items inspired by this thesis:<ul><li>“I expect surprising recommendations” vs. “I actually receive surprising recommendations”</li><li>Rate satisfaction with novelty, diversity, and unexpectedness</li></ul></li><li>Segment results by platform behavior, engagement style, and demographics</li></ul></td></tr>
<tr><td>Monitor trust perception</td><td><ul><li>Implement in-platform feedback such as “Do you trust our recommendations?” or “Do you feel in control of your suggestions?”</li><li>Implement in-suggestion feedback such as a thumbs up or down next to the suggestion heading</li><li>Monitor implicit signals and user behavior in conjunction with feedback (how often users use the control options)</li></ul></td></tr>
</table>

| | |
|---|---|
| **Monitor curiosity** | • Measure using explicit settings (if available, e.g., "Choose your recommendation style") or implicit signals (e.g., how often user sways toward unknown vs. similar content)<br>• Identify clusters of high curiosity vs. passive users |
| **Identify serendipity-sensitive users** | • Use the above data to define who is most receptive to surprise and which contexts enable them (e.g., device type, time of day, location)<br>• Look for behavior patterns between users |
| **Phase 3. Design for Trust and Control** | |
| **Set discovery expectations** | • Draw attention to discovery options on signup walkthrough<br>• Reiterate platform capabilities after feature updates |
| **Build and signal trust** | • Use rationale explanations (e.g., "Because you liked X…")<br>• Include external validation (e.g., "Critically acclaimed", "Popular among users like you") for new additions<br>• Add popularity signals (views, likes, ratings) where applicable |
| **Introduce or enhance control features** | • Draw attention to where users can remove or exclude content from their taste profile<br>• Offer session-specific opt-outs (e.g., "Don't track this session")<br>• Increase transparency of taste profiles (provide overview or visualizations) |
| **Tailor to segments** | • Push more exploratory content to serendipity-receptive users and measure impact<br>• Use micro-messages for more sensitive users (e.g., "You've browsed items like this, want to try something new?") |
| **Phase 4. Engineer Adaptive Serendipity** | |
| **Expand beyond-accuracy metrics** | • Incorporate novelty, diversity, and serendipity into model evaluation<br>• Use both offline and online (A/B) testing |
| **Deploy exploration-enhancing techniques** | • Multi-armed bandits to balance relevance and discovery<br>• Hybrid human-AI curation<br>• Context-aware models using session-based cues or external context |
| **Label serendipitous content** | • Add cues like "Hidden gem", "You might not expect this", or "Out of your usual pick"<br>• A/B test to evaluate which labels improve trust and engagement |
| **Create customizable discovery modes** | • Implement toggles or pages:<br>    o "Safe" with most familiar content and less novelty and unexpectedness<br>    o "Balanced" with a standard mix<br>    o "Surprise me" with increased serendipity (novelty, diversity, and unexpectedness together)<br>• Track user selections and evolve defaults based on behavioral feedback |
| **Phase 5. Monitor and Iterate** | |
| **Implement micro-feedback systems** | • Measure serendipity satisfaction with "Was this a good surprise?" or "Want more or less like this?"<br>• Train models using this input as a reward signal |
| **A/B test adaptive serendipity** | • Test algorithm and suggestion variations on user segments<br>• Track metrics like click-through rates on discovery content, session length, drop-off rate, and diversity consumed<br>• Combine behavioral metrics with serendipity micro-feedback (above) to train model to adapt serendipity level to each user over time |
| **Run iterative feedback loops** | • Re-run expectation-experience surveys quarterly to track how platform is closing or widening the gaps<br>• Optimize measurement of trust, curiosity, and control perception as new features roll out and user frequency increases |