Course code: SKDK11
Term: Spring 2024
Supervisor: Ilkin Mehrabov
Examiner: Nils Holmberg

Balancing Personalization and Privacy

A quantitative study examining the factors influencing the willingness to receive personalized communication on streaming platforms.

WILMA DAHLIN & LINA HJORTSMARK

Lund University
Department of Strategic Communication
Bachelor´s Thesis

A big thank you…

To our supervisor Ilkin Mehrabov for your expertise and patience with our many questions during the last couple of months.

To the Department of Strategic Communication for equipping us with invaluable knowledge over the past three years, preparing us for the future.

To our classmates for providing us with valuable advice not only during this thesis period, but throughout the entire bachelor program.

To each other for our effective and rewarding collaboration during this thesis period. We would like to underline that we have both contributed equally to the study.

Wilma Dahlin

Lina Hjortsmark

Lund

May 2024

Abstract

Balancing Personalization and Privacy

A quantitative experimental study examining the factors influencing the willingness to receive personalized communication on streaming platforms.

The aim of this study was to examine which factors influence the willingness to receive personalized communication based on personal information on streaming services. This was achieved using the Privacy Calculus Model (PCM), which includes the factors of perceived benefits and risks. In addition to investigating the significance of the original PCM parameters, the intention was to reveal whether the significance of the factors trust, transparency, and context, could impact the willingness to receive personalized communication based on personal information. To examine the possible influence of context, the streaming services Netflix and Spotify were used as examples to illustrate possible differences in attitudes towards foreign and domestic companies. The study employed a quantitative experimental design, and the empirical data was collected through two online surveys targeting young people from Sweden through a convenience sampling method. Given the experimental approach, the questionnaires were randomly distributed among respondents, resulting in 95 usable responses per survey. To analyze the empirical data, a Multiple Regression Analysis was performed using SPSS. The findings indicated that trust had the most significant impact on the willingness to receive personalized communication for Netflix and trust in combination with transparency had the most significant impact for Spotify. These findings can contribute to valuable insights for communicators applying personalization strategies in their work. The study suggests that companies should prioritize strategies aimed at building trust and transparency to balance the complex dynamics between personalization and privacy.

Keywords: Personalized communication, Personalization Privacy Paradox, Privacy Calculus Model, Benefits, Risks, Trust, Transparency, Context, Streaming services, Netflix, Spotify.

Number of characters, including spaces: 87 777

Table of contents

1. Introduction

The first section of our thesis presents the background to the study topic, followed by the problem definition, purpose and research question. Furthermore, the study's research focus, delimitations and the thesis outline are presented.

1.1 Background

The digital landscape is constantly evolving, requiring businesses to adapt to new technologies to meet customer needs. The changing media landscape has also resulted in an increased amount of information available to consumers, creating a variety of impressions every time they use a digital service. This has put pressure on companies to stand out from the crowd, and many companies have shifted their focus from mass marketing to targeted marketing as this has proven to be a successful strategy to capture the attention of today's consumers (Palmatier et al., 2018). Consequently, this has created a relatively new trend in digital marketing, personalized communication, where marketers are tailoring their messages to individuals rather than groups of people. According to the Twilio Segment (2023) report, 56 percent of customers claim that a personalized experience will lead them to become repeat buyers, an increase by 7 percent from 2022. Additionally, 80 percent of company leaders report that consumers spend 38 percent more money when their service experience is personalized. At the same time, today’s consumers desire more control over their personal information and are increasingly aware of how their data is being used (Twilio, 2023).

Falkheimer and Heide (2018) argue that strategic communication is a product of late modernity. Furthermore, they explain that societal trends have a central role in shaping the future of strategic communication. Communication will always be important, thus it is essential to keep up with the times to know how to best deliver it to achieve organizational goals. A contemporary trend that is very fundamental is individualization. People increasingly expect communication from companies to be personalized for each individual. This is made possible by the new media technologies that are constantly evolving. The ability to learn more about stakeholders through their digital data is a great opportunity for strategic communication, since customizing messages to a specific target group is one of the key elements of strategic communication (Falkheimer & Heide, 2018).

One of the industries that has become more popular in the era of digitalization is streaming services (Tsigkari et al., 2024). These companies have revolutionized contemporary media consumption habits by enabling individuals to stream a diverse array of movies, music, and TV series on demand, obviating the need for downloading content (Bender et al., 2022). Moreover, it is an industry that has adopted advanced personalization strategies. Streaming services today use personalization through algorithms and recommendation systems where they target their content depending on their customers' individual preferences (Tsigkari et al., 2024). However, as personalization requires personal data, it also comes with privacy concerns for the customers. Therefore, it is important for companies to know what level of personalization to apply in order to create a balance between personalization and privacy (Falkheimer & Heide, 2018). Thus, privacy concerns linked to personalization is an important subject to explore, in order to develop the field of strategic communication.

1.2 Problem definition

The increased amount of personalized communication has resulted in people today expecting communication from companies to be tailored to them (Alimamy & Gnoth, 2022). One context where personalized communication and tailored offers are common is in streaming services (Tsigkari et al., 2024). Personalization aims to provide people with an enhanced user experience by delivering content that is aligned with individual preferences and behaviors, thereby increasing satisfaction (Hayes et al., 2021). However, as personalized communication has increased, so have people's privacy concerns (Wang et al., 2024).

People are worried that their personal data will be misused online and that their privacy will be compromised. This has created a paradox between people's desire for personalized communication and their concerns about privacy, known as the personalization-privacy paradox (Cloarec et al., 2024b). Given that personalized communication predominantly relies on individuals' personal information, it poses a significant challenge for companies when individuals are unwilling to disclose such information. Thus, if individuals consider the privacy concerns to outweigh the benefits of personalized communication, companies face the risk of encountering greater difficulty in reaching these people (Hayes et al., 2021). This tension is therefore something that marketers and companies must balance in order to enhance consumer satisfaction and loyalty.

Previous research on personalized communication underscores the significance of comprehending how individuals assess the perceived benefits of personalization against the perceived risks of disclosing their personal information online (Wang et al., 2024). This assessment, known as the Privacy Calculus Model (PCM), can be seen as part of the paradox mentioned above. The growing desire for personalized communication combined with growing privacy concerns makes this model interesting to further examine. While previous research has highlighted general principles of the PCM, there is a research gap when it comes to applying the model in the context of streaming services. Examining the PCM through streaming services is important given the widespread use of these platforms (Internetstiftelsen, 2022). Streaming services often rely heavily on user data to make content recommendations and deliver personalized experiences (Tsigkari et al., 2024), making them an optimal setting to explore the PCM.

Additionally, there is limited research examining whether the nature of the company handling individuals' personal information influences their level of privacy concerns. For instance, people may consider the context of where the company operates in their decision of disclosing personal information. Moreover, the domestic or foreign status of a company can shape the level of trust customers place in it, which consequently could affect their attitudes towards personalized communication (Bhattacharya et al., 2023).

Furthermore, it is crucial to continue investigating the PCM since there may be additional factors beyond the perceived benefits and risks that influence people's willingness to disclose personal information. In a study conducted by Bile Hassan et al. (2022), the researchers explored the PCM in combination with the Unified Theory of Acceptance and Use of Technology (UTAUT2). Their developed model could be validated through their research. Another study by Chen (2018) examines the PCM in an Asian social media context. The researcher extended the model with factors such as social capital, privacy self-efficacy and privacy management as they were assumed to have an impact on individuals' intention to share personal information. These studies demonstrate the importance of continuing to revisit the PCM in different settings to understand the complexity of balancing privacy and personalized communication.

1.3 Aim and research question

The aim of this thesis is to examine which factors influence the willingness to receive personalized communication on streaming services. Furthermore, the study aims to investigate the intention to disclose personal information, as it is a prerequisite for personalized communication. To accomplish this, the Privacy Calculus Model (PCM) will be used to examine how perceived benefits and risks influence the intention to disclose personal information. According to previous literature, additional factors can influence an individual's intention to disclose personal information. Therefore, the factors trust, transparency and context will be included to expand the PCM. To examine whether the context has an impact, the American streaming service company Netflix will be compared with the Swedish streaming service company Spotify. The study adopts a Swedish perspective, therefore these companies are relevant as they are among the two most used streaming services in Sweden (Internetstiftelsen, 2022). Additionally, they are both at the forefront in terms of their personalization strategies and advanced recommendation systems (Tsigkari et al., 2024). Through a quantitative experimental method, the study aims to investigate which of the factors above mainly influences streaming service users' willingness to receive personalized communication based on personal information. To address these issues, the following research question has been formulated:

> **RQ:** *Which factor of perceived benefits, perceived risks, trust, transparency and context have the most impact on streaming service users' willingness to receive personalized communication based on personal information?*

1.4 Research focus

As previously mentioned, the streaming service industry is at the forefront of personalization (Tsigkari et al., 2024). Hence, the research focus of this study will be Netflix and Spotify, two of the most utilized streaming services in Sweden (Internetstiftelsen, 2022), aiming to explore whether their origin country significantly impacts the research outcomes. The following section provides a brief explanation of the two companies, and how they collect and utilize their users' personal data to personalize their services.

1.4.1 Netflix

Netflix is an American company that offers a wide range of TV series and films (Netflix, n.d.b). Netflix was founded in 1997 with a business idea of renting out DVDs by mail in the United States. In 2000, they launched a recommendation system based on customer data to predict future customer choices. The company continued to evolve and in 2007 they began offering their customers to stream TV series and films via the internet. Their service began to spread beyond the United States, and in 2012 it became available in the Nordic countries. Today, Netflix has over 200 million members worldwide (Netflix, n.d.a) and is a subscription-based service that offers different features depending on the subscription users pay for (Netflix, n.d.b).

Netflix collects a large amount of data from its users. This includes information that users themselves provide to Netflix, such as email address, phone numbers, payment details and content preferences. The information is collected partly when a user registers with Netflix, but also in other ways such as when contacting customer service. Netflix also collects information automatically, including all the different choices a user makes when utilizing the service. For example, how users react to notifications and text messages from Netflix, their network devices, streaming devices, IP addresses and much more. Additionally, Netflix collects data from their partners and other sources. The purpose of this is to comprehend their audience to optimize their service to match the preferences of each individual person (Netflix, 2022).

1.4.2 Spotify

Spotify is a Swedish company launched in 2008. It is an audio streaming service that has gathered millions of songs and podcasts on its platform. Spotify's business model is to offer their customers to stream music without downloading or buying it. Today, Spotify is the largest subscription-based audio streaming service in the world, with 236 million subscribers. However, Spotify also offers its service to non-payers and therefore has over 602 million users (Spotify, n.d.).

Spotify collects personal data from its users in order to improve the algorithm and offer personalized and tailored content to each user. The type of data collected depends on various factors, such as country and subscription type. Usually, Spotify collects information including profile name, password, email address, phone number, country, street address, date of birth, gender and payment details. Some of this data is provided by the user upon registration, and other data is collected through the user's device. Additionally, Spotify collects information about the user’s behavior on the platform, technical data, voice data, motion and direction-generated device sensor data and data from third parties (Spotify, 2023).

1.5 Delimitations

This study has been limited to the factors perceived benefit, perceived risk, trust, transparency and context, assuming to influence the willingness to receive personalized communication based on personal information. We are aware that there may be more factors that affect the willingness to receive personalized communication. However, these factors are considered most relevant in this study according to previous research, which will be described in detail in 2.4. Thus, while our focus is limited, we are open to the complexity of the subject.

The sample has been limited to people in Sweden aged 20-30 who use the streaming services Netflix or Spotify. By focusing on this specific demographic group, we aim to capture insights relevant to the Swedish context, and preferences of young people. One reason for this delimitation was that this age group is some of the most frequent users of streaming services (Internetstiftelsen, 2022). Moreover, given that convenience sampling was employed as the study’s sampling method, it was considered appropriate to select a population easily accessible to us as we belong to this demographic group.

Finally, we decided to examine the willingness to receive personalized communication based on personal information from a user perspective only. By prioritizing the user perspective, our goal is to provide valuable insights to communicators. Understanding users' attitudes towards personalized communication can provide guidance in the design of communication strategies, to ensure that personalization does not compromise anyone's privacy.

1.6 Thesis outline

The structure of this thesis is organized as follows: First, we introduce an overview of the previous literature within digital marketing communication followed by an outline of personalized communication and recommendation systems. Secondly, we describe the personalization privacy paradox and the privacy calculus model on which our research model is based. Moving forward, we proceed to develop our hypotheses and present the constructed research model of the study. This is followed by our methodology section, which provides the study’s scientific approach, research design, sampling strategy, analysis method, limitations and ethical considerations. We then present the obtained results of the study and conclude by testing whether our posed hypotheses are supported. Moving forward, the results of the study are discussed in relation to previous research, followed by a conclusion that answers the study's research question. Lastly, we address the limitations of the study and outline suggestions for future research.

2. Literature review

The following chapter will present an overview of previous literature. Given that this study delves into personalized communication, a strategy in digital marketing communication, the chapter starts with a brief introduction to the subject. Furthermore, theories connected to our hypotheses development and research model are presented.

2.1 Digital Marketing Communication

Fierro et al. (2017) explain that with the advent of the internet and the development of technology, the landscape of marketing communications changed. This initiated a new era in which marketers had to adapt to the new technology, leading to the emergence of digital marketing communication. Digital marketing communication refers to all marketing that takes place on digital platforms. It was used for the first time in the 1990s (Fierro et al., 2017). However, it was not until the 2000s and 2010s, that it really took off and became more refined (Fierro et al., 2017; Sharma & Aggarwal, 2023). Unlike traditional marketing, digital marketing enabled an easy way to measure the results of the advertisements. Through these measurements, marketers gained a more precise understanding of their customers' interests. The increased use of digital media also enabled the collection of user data, such as which platforms they visit, their search history and what they consume online. All this data is stored, creating a digital footprint. When combined with advertising metrics, this footprint enabled marketers to target their consumers more effectively (Mehta & Kulkarni, 2020).

2.2 Personalized Communication

One strategy within digital marketing communication is personalization. This concept has the aim to "deliver the right content to the right person at the right time" (Aguirre et al., 2015, p. 35). This means that companies create consumer profiles based on their customers personal data in order to target them with relevant services or products. Personalized communication often appears in two forms, personalized advertising and personalized services (e.g. recommendations) (Karwatzki et al., 2017), where our study refers to it as personalized services. When formulating a personalized message, marketers begin with the consumer to design the advertisement according to the consumer's preferences (Mehta & Kulkarni, 2020).

This is done by analyzing the customer's personal data collected from their online behavior (Hayes et al., 2021). Through personalization, a message or service is tailored to meet and fulfill the interests of individual users. The goal of this is to increase the user's willingness to consume (Ameen et al., 2022). In fact, personalization is often crucial to the success of digital marketing, as consumers rarely look at an online advertisement unless it is relevant to them (Aguirre et al., 2016).

Personalized communication comes with several benefits for both marketers and consumers. By tailoring messages to specific individuals, marketers can target customers they know are potentially profitable. Thus, it can be a time and cost-effective strategy since it avoids spending resources on people who will have no interest in what is being promoted (Hayes et al., 2021). Pukas (2022) explains that personalization tends to improve customer loyalty and satisfaction. Moreover, personalization strategies help companies create stronger relationships with the customer (Behera et al., 2020; McKee et al., 2023). However, as personalization relies on a large amount of consumer data to be successful, it comes with potential risks regarding consumers' privacy online (Cloarec et al., 2024a).

2.2.1 Recommendation Systems

One part of personalized communication is recommendation systems (Behera et al., 2020). It is a type of information filtering technique that takes advantage of the user's personal information to predict and recommend products that the user is likely to enjoy (Yildiz et al., 2023). The increasing use of digital media results in greater exposure to a vast amount of information that can be difficult to sort. Recommendation systems are therefore essential to reduce information overload by filtering out what is not relevant to the consumer (Wang et al., 2024; Yildiz et al., 2023).

Recommendation systems are central to both Netflix's and Spotify's services (Khoo, 2023; Tofalvy & Koltai, 2023). Their recommendation system is a combination of collaborative filtering and content-based filtering (Khoo, 2023; Anthony et al., 2022). Content-based filtering is a strategy that provides the user with content that is similar to what the user has previously liked or interacted with (Bobadilla & Gutiérrez, 2024). Collaborative filtering is a strategy in which users' past choices and behaviors are compared with other similar users. Through this comparison, collaborative filtering generates predictions and recommendations for future content that align with the user’s preferences and interests (Khoo, 2023; Behera & Nain, 2023).

2.3 Personalization-Privacy Paradox

Although personalization may seem beneficial, there are also disadvantages from a consumer perspective. Since personalized communication depends on consumer data, individuals have grown increasingly apprehensive about their online privacy. As a result of this, they can be reluctant to disclose personal information online (Wang et al., 2024; Cloarec et al., 2024a). However, Wang et al. (2024) explains that even if people are concerned about their privacy, they still want personalized advertising and offers. This creates a tension between the desire to withhold personal information and the desire to receive personalized communication. This is explained as the personalization-privacy paradox (Kronemann et al., 2023; Wang et al., 2024). Thus, people want the tailored experience that personalization offers, yet they are worried about sharing their personal data because there is a fear that it will be misused by companies (Wang et al., 2024).

2.3.1 Privacy Calculus Model

Despite people's privacy concerns, it is common that they disclose their personal data or agree to its use online if it can bring some kind of benefit. This phenomenon can be explained by the Privacy Calculus Model (PCM), which can be seen as part of the personalization-privacy paradox. The PCM indicates that people consider the act of disclosing their personal data by comparing the benefits and risks of doing so (Wang et al., 2024). If consumers consider the risk of disclosing personal information to be greater than the benefits that may come with it, they are less willing to disclose the information. On the other hand, if the benefits are expected to outweigh the risks, consumers will feel more comfortable disclosing personal information (Hayes et al., 2021; Tang & Ning, 2023). Hence, the PCM explains how advantages and disadvantages can be set against each other to consider what is more important (McKee et al., 2023). Wang et al. (2024) and Cloarec et al. (2024b) explain the PCM as a way for people to analyze what they should do to maximize their returns, and the consideration of benefits and risk play a major role in the attitude people will have towards personalization.

2.4 Hypothesis development

Previous research indicates that the Privacy Calculus Model is an important part of understanding people's attitudes and behavioral intentions towards personalized communication online (Wang et al., 2024). Hence, our study will be based on the key principles of the PCM. To further understand which parameters influence the willingness or unwillingness to receive personalized recommendations, the study’s model has been developed. This decision is based on previous research that underlines that there may be more aspects than just benefits and risks that affect the intention to disclose personal information. Thus, the following three factors have been added to the model: trust, transparency and context. The reason for this choice is explained further in section 2.4.3, 2.4.4 and 2.4.5.

2.4.1 Benefits

According to the PCM, people weigh the benefits and risks of sharing personal information to determine which is more important before making a decision (Wang et al., 2024). Bol et al. (2018) suggested that the primary motivation behind people's intention to online self-disclosure is the expected benefits from doing it. The perceived benefits can be hedonic or utilitarian and include for example social support, entertainment, rewards and tailored information (Bol et al., 2018; Wang et al., 2024). The more benefits people expect, the more likely they are to disclose their personal information and vice versa (Bol et al., 2018).

In a study by Shim and Yeon (2022) it is explained that when people perceive that the benefits of sharing personal information outweigh the privacy risks involved, they sacrifice a certain degree of privacy in exchange for benefits. An individual may use a streaming service for different purposes, possibly to fulfill some kind of psychological, social or personal need. By utilizing the service, the user obtains various types of gratification. Hence, the PCM is applicable as streaming service users will receive benefits in terms of preferred content, in exchange for their privacy (Shim & Yeon, 2022).

This study includes the benefits of ease of use, content satisfaction and time savings as these are most appropriate in relation to streaming services. The aim is to measure whether these specific benefits make it worth sacrificing personal information. Hence, the study’s first hypothesis is:

H1: *Perceived benefits will have a positive impact on the willingness to receive personalized communication based on personal information.*

2.4.2 Risks

Perceived risk is about the potential negative outcomes that can happen if one shares personal information (Mohammed & Rozsa, 2024). Hayes et al. (2021) explain that if people consider the risks of sharing personal information to be greater than the benefits of doing so, they will be less willing to disclose the information. Previous research highlights several issues that can be seen as risks of sharing personal information online. Jiang et al. (2022) explain that a central part of the perceived risks are privacy concerns. Furthermore, the authors explain that this may involve a concern that the act of sharing personal data leads to negative consequences and a perceived loss of control over one's data.

Another perceived risk is getting caught in a filter bubble. Gong et al. (2024) describes filter bubbles as individuals being restricted within an algorithm and presented exclusively with content aligned with their preferences. Furthermore, Gong et al. (2024) explain that this can be seen as a positive part of personalization, but the risk lies in being limited to content that continuously reinforces existing beliefs, leading to minimal exposure to diversity. Individuals may however find it difficult to acknowledge that they are in a filter bubble, and thus potentially overlook its consequences (Gong et al., 2024).

According to Plangger and Montecchi (2020), a further risk that people may experience when sharing personal information with companies for the purpose of personalized communication is the feeling of surveillance. Similarly, Gironda and Korgaonkar (2018) explain that personalized recommendations that are considered excessively tailored can induce feelings of discomfort and intrusiveness. Furthermore, they argue that if recommendations feel intrusive, individuals are less likely to share more personal information. Thus, based on above research, the second hypothesis of the study is:

> ***H2:*** *Perceived risks will have a negative impact on the willingness to receive personalized communication based on personal information.*

2.4.3 Trust

Trust refers to the reliance on different types of partners and is based on the belief that they are reliable (Mohammed & Rozsa, 2024). Bleier and Eisenbeiss (2015) explain that trust is an important element in how people perceive companies' marketing and recommendation systems. Moreover, the authors claim that greater trust in a company correlates with a reduced concern about hidden motives behind their recommendations, such as only aiming to maximize the company's profits.

Bleier and Eisenbeiss (2015), and Cloarec et al. (2024b) argue that personalized recommendations can be intimidating for people, as they are based on a large amount of personal data. Furthermore, Cloarec et al. (2024b) explain that this leads to an increased need for trust as the disclosure of personal information can create a sense of vulnerability. However, depending on the level of trust one has in the company from which the recommendations come, the sense of intimidation can decrease or increase (Bleier & Eisenbeiss, 2015). If people have a high level of trust in the company, they are more willing to accept recommendations based on their personal information. On the other hand, when trust levels are low, there is a greater risk of experiencing privacy concerns and people are more critical of what is recommended (Bleier & Eisenbeiss, 2015; Cloarec et al., 2024b).

Thus, trust plays a major role in how personalized recommendations will be perceived. Hayes et al. (2021) also argue that trust is a significant factor in the willingness to disclose personal information to companies. Thus, previous research indicates that trust may have an influence on this study's dependent variable, which makes it a relevant factor to include and examine. Hence, the third hypothesis of the study has been formulated as follows:

H3*: Trust in a company will have a positive impact on the willingness to receive personalized communication based on personal information.*

2.4.4 Transparency

Li et al. (2023) argue that it is important for companies to be transparent in their communication about how they collect and use their customers' data. Through transparent communication, customers have an opportunity to make informed decisions about whether or not to disclose their personal data (Li et al., 2023; Cloarec et al., 2024a). A study by Betzing et al. (2020) shows that although companies in Europe are legally required through the General Data Protection Regulation (GDPR) to communicate how they process personal data, for example through privacy policies, few service users actually read or understand them. Furthermore, Betzing et al. (2020) explain that enhancing the transparency of data processing practices increases users' understanding of their consent decisions. Increased transparency is therefore essential to enable service users to make informed decisions about data protection (Betzing et al., 2020).

Karwatzki et al. (2017) emphasize that it is important to distinguish between privacy policies and transparency features. While privacy policies provide a comprehensive description of the company’s data protection policy, transparency features provide an overview that improve the understanding of what information is collected and used. Transparency features are presented in a way that is accessible and understandable to consumers, unlike many privacy policies. Karwatzki et al. (2017) study indicates that when privacy information and data protection is clear and easily accessible, consumers are more likely to disclose their personal information. Moreover, the Twilio (2023) report underlines the importance of transparency to address customers' growing privacy concerns. Companies can address customers' privacy concerns by providing clear communication of how their data is used, thus creating a stronger foundation for successful personalization efforts (Twilio, 2023).

Karwatzki et al. (2017) and Cloarec et al. (2024a) highlight that transparency is an important factor in the level of trust people have in a company, and that high transparency leads to increased trust. Furthermore, Kawaf et al. (2023) emphasizes that transparency about how data is used is crucial for people to be willing to share their personal data. Similarly, Cloarec et al. (2024a) suggests that a high level of transparency can lead to a greater acceptance of information sharing. Hence, this study's fourth hypothesis is:

H4: *Transparency will have a positive impact on the willingness to receive personalized communication based on personal information.*

2.4.5 Context

In a study where Hirschprung (2023) examined the privacy paradox in different domains, the author concluded that the privacy paradox varies depending on the domain. Thus, people's privacy concerns and their actual behavior regarding their personal data were different depending on the context (Hirschprung, 2023). Similarly, Aguirre et al. (2016) argue that people's privacy concerns often are context-based. Canhoto et al. (2023) reinforce this by highlighting that the context in which personalized communication takes place, has an impact on the attitude people will have towards it.

Augirre et al. (2016) give the example that when the context is a well-known company, people are more comfortable with sharing their personal data. As Netflix and Spotify are among the most widely used streaming services in Sweden (Internetstiftelsen, 2022), one can assume that there is not much difference in how well known the two services are. However, Wang et al. (2023) also explain that cultural differences play a major role in how privacy concerns are perceived and what is considered private and not. The cultural differences in this study is that Spotify's country of origin (COO) is Sweden and Netflix's is the US. Since the study only examines the opinions of Swedes, Spotify has a national perspective and Netflix an international one. In a study by Bhattacharya et al. (2023), the authors claim that COO influences not only the consumers' decision-making and how they perceive companies, but also their perception of trust and privacy. Furthermore, the authors explain that in many cases there is home country bias, and that people tend to prefer companies from their own home country over foreign companies, as they trust those from their home country more. Hence, the studies fifth hypothesis is:

> ***H5:*** *The context will have a significant impact on the willingness to receive personalized communication based on personal information.*

2.5 Research model

Based on the theories and previous research discussed above, five hypotheses have been formulated that are expected to have a significant impact on the willingness to receive personalized communication based on personal information. All five hypotheses are listed below:

> ***H1:*** *Perceived benefits will have a positive impact on the willingness to receive personalized communication based on personal information.*

> ***H2:*** *Perceived risks will have a negative impact on the willingness to receive personalized communication based on personal information.*

> ***H3****: Trust in a company will have a positive impact on the willingness to receive personalized communication based on personal information.*

> ***H4:*** *Transparency will have a positive impact on the willingness to receive personalized communication based on personal information.*

H5: *The context will have a significant impact on the willingness to receive personalized communication based on personal information.*

The hypotheses of the study are included in the following model as independent variables showing their expected relationship with the dependent variable:

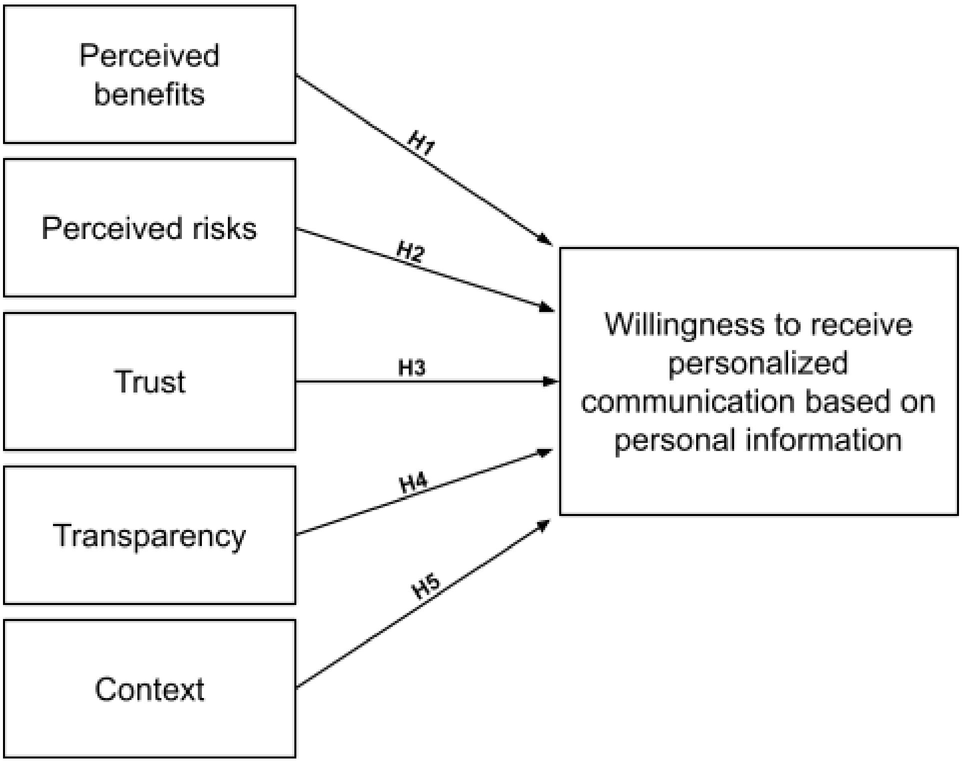


Figure 1: *The extended Privacy Calculus Model with the additional factors: trust, transparency and context.*

3. Methodology

The following chapter will present the scientific approach of the study, followed by research design, sampling strategy and method of analysis. Lastly, the study's methodological limitations and ethical considerations are discussed.

3.1 Scientific approach

This study adopts positivism as its epistemological position. This means that only phenomena that can be confirmed by objective empirical observations can be considered as knowledge. The positivistic approach states that theories are only valuable if they can be confirmed by observations. Thus, the study adopts a deductive approach, meaning that theory and previous research are used to generate hypotheses that can explain patterns of reality. Furthermore, since knowledge is reached through objective observations, the ontological approach of this study is objectivism. This means that we define reality as existing separately from interpretations and social constructions. These scientific approaches are relevant to the research method of this study, as objective data is collected through a survey where we remain neutral (Clark et al., 2021).

3.2 Research design

This thesis is based on small-scale quantitative research, meaning that it aims to draw measurable inferences that can be generalized to a similar context outside its sample. The quantitative approach was applicable as the data would be statistically analyzed to determine whether the study's hypotheses were supported or rejected (Pallant, 2020). To examine the impact of the hypotheses on user’s perceptions towards personalized communication, an experimental design was applied. The experimental design involved manipulating one of our independent variables to determine its impact on our dependent variable. In the case of this study, the independent variable being manipulated, i.e. the stimulus, was the context. Thus, respondents were to answer a questionnaire about either Spotify or Netflix. By manipulating the context, we were able to examine whether it had any causal relationship with our dependent variable (Boyle & Schmierbach, 2024).

Furthermore, the study adopted a between-subjects approach. This means that the participants were randomly assigned to one of the two conditions in order to compare the results between two groups (Boyle & Schmierbach, 2024). Respondents were randomly assigned to a group, either Sample 1: Netflix, or Sample 2: Spotify. Hence, Sample 1 was exposed to the Netflix condition and Sample 2 was exposed to the Spotify condition. By using an experimental design, the results could be compared to measure any differences in respondents' answers (Clark et al., 2021). To conduct the experiment, two online surveys were created. Moreover, the surveys were constructed with a cross-sectional design, meaning that the data was collected at a single point of time on a sample of cases (Clark et al., 2021). The questionnaires were constructed with identical questions and structure, with the only variation being the words Netflix or Spotify in the contextual questions. The items of the questionnaires were created on the concepts of the Privacy Calculus model, as well as the additional variables that are based on previous research, as presented in section 2.4.

3.2.1 Survey design

As mentioned above, the chosen data collection method for the study was online surveys. Trost and Hultåker (2016) argue that surveys are suitable when the purpose is to investigate people's knowledge, attitudes and behaviors. Therefore, this method was considered appropriate for the purpose to examine people's willingness to receive personalized communication based on personal information. The surveys were self-administered, meaning that participants answered the questionnaire themselves without the presence or assistance of us. This required that questions, descriptions and instructions were clearly formulated in order to avoid misunderstandings by respondents that may lead to incorrect results (Boyle & Schmierbach, 2024).

The surveys were constructed using the tool Google Forms. They were sent out on the 15th of April 2024 on our private social media channels, and closed after one week, on the 23rd of April 2024. The advantage of creating online surveys in Google Forms instead of postal surveys is that we were able to be more flexible and include interactive elements, for example filtering options (Trost & Hultåker, 2016). The study benefited from this, since we were able to sort out the respondents who did not belong to the population of people in the age of 20-30, who are users of Netflix or Spotify and Swedish citizens. Another advantage of online surveys is that one can ensure the anonymity of the respondents, which could contribute to a higher response rate and honest answers (Boyle & Schmierbach, 2024).

The questionnaires started with an introduction where participants were given information about the purpose of the study and how their responses would be processed ethically. Additionally, they were provided with our definition of personalized communication and personal information to ensure the participants shared our understanding, increasing the validity of the study (Bryman, 2012). The surveys were divided into seven parts with a total of 32 items. The first section dealt with demographic factors (age, gender, level of education and occupation). The questions regarding gender and occupation were measured through a nominal scale, educational level was measured through an ordinal scale and age was measured through a ratio scale. The remaining parts included questions (see appendix 7.1) that measured all of the variables presented in section 2.4: benefits, risks, trust, transparency and context. These questions were formulated as statements with a five-point Likert scale, from “strongly disagree” to “strongly agree”, as response options. The purpose of this is for respondents to answer to what extent they agree with the statement. We decided to use a Likert scale from one to five since more numbers would have been irrelevant to our statements, and may make it difficult for respondents to answer (Boyle & Schmierbach, 2024). One advantage of using a Likert scale is that it facilitates the process of combining items that represent the same concept to create indexes (Boyle & Schmierbach, 2024). Furthermore, the Likert scale provides a better estimate of the respondents’ opinions in a nuanced way and is suitable when performing a statistical analysis (Pallant, 2020).

In order to implement the experiment and randomly distribute the questionnaires to every other respondent, the website allocate.monster was used. This tool allowed us to merge the two different survey links into one single link that opened every other survey when respondents clicked on it (Fergusson, 2016). Thus, respondents had an equal chance of being assigned to the Netflix questionnaire as to the Spotify questionnaire. By randomly distributing the conditions, the validity of the study could be increased (Boyle & Schmierbach, 2024).

3.3 Sampling strategy

The population for this study could theoretically have been extensive, as Netflix and Spotify are used by a broad population (Internetstiftelsen, 2022). However, due to limited time and cost resources, the study's population has been limited to Swedish 20-30-year olds. The reason for choosing this population is because it is similar to our demographic characteristics.

Thus, this population was considered the easiest to reach in order to get a reasonable normal distribution in the collected data. Furthermore, this was a relevant population as people in this age group are some of the most frequent users of streaming services (Internetstiftelsen, 2022). Due to the resources and the identified population, a convenience sampling strategy was considered suitable. A convenience sampling strategy is a non-probability sample where individuals are selected only based on their availability (Boyle & Schmierbach, 2024).

Convenience samples cannot be generalized to a larger population since the individuals are not randomly selected on equal terms. Hence, the results of convenience sample studies tend to contribute less scientific value than random sampling strategies (Clark et al., 2021). Therefore, it is necessary to ensure that the sample is of an acceptable size in order to be of any scientific value. Pallant (2020) explains that a formula to calculate an appropriate sample size is $N > 50+8m$ (m is the number of independent variables). Since this study has five independent variables, the formula in this case is as follows, $N > 50+(8x5) =90$. This means that this study needed at least 90 respondents per survey to be of scientific value. In order to reach the population and achieve the approved sample size, data collection was done by distributing the survey on our private LinkedIn, Instagram and Facebook accounts.

Despite this calculation, we are aware of and want to underline the shortcomings of convenience sampling. Generalizing our findings is challenging with this strategy, as the sample is probably not representative to our population. This creates a selection bias and reduces external validity (Clark et al., 2021). To improve the validity of the study somewhat, an experimental method was chosen where respondents were randomly distributed to a survey, as described in section 3.2. This helps to create a small additional level of randomness to compensate for the shortcomings with a convenience sample. Even if convenience sampling cannot be considered as an ideal method, it is frequently used by social researchers and can be acceptable if the researcher is aware of its limitations (Clark et al., 2021).

3.3.1 Pilot study

Before sending out the final questionnaires, a small-scale pilot study was conducted on the 12th of April 2024. This was done to ensure that the design of the questionnaires was reasonable and the questions were understandable. By doing this, the questionnaires could be tested to assess if anything needed to be changed before it was sent out to the study's population (Wenemark, 2023).

Wrench et al. (2013) state that for a pilot study to be valuable, it should consist of 5-10 percent of the study's population. As previously mentioned, this study requires 90 respondents for each survey in order to be of scientific value. Thus, five people completed each pilot survey, resulting in us reaching the suggested percent (5 out of 90 respondents equals 5.5 percent). The participants' suggestions for improvement helped us to correct shortcomings in our measurement instrument, such as syntax errors. The purpose of this was to enhance the validity of our study (Boyle & Schmierbach, 2020). However, even if the proposed percentage was achieved, the number of participants remains low. Consequently, certainty regarding the increase in validity cannot be established.

3.4 Method of analysis

When the data collection was completed, all survey responses were transferred to an Excel file that was inserted into the Statistical Package for the Social Sciences (SPSS) to begin the analysis. A total of 96 people participated in the Netflix survey and 95 people participated in the Spotify survey. To prevent outliers in the data, the surveys were designed with demographic filter questions to ensure that responses were only collected from people in our sample. However, the item regarding if the respondent was a Netflix or Spotify user was not designed as a filter question. Therefore, the collected data were carefully reviewed to identify potential outliers in that aspect.

The study’s selected method of analysis was a Multiple Regression Analysis (MRA). This method is useful to examine to what extent the variance in the dependent variable is explained by the independent variables (Pallant, 2020). In our case, this means that the MRA determined which of the five independent variables had the most significant impact on the willingness to receive personalized communication based on personal information. The logic behind MRA is related to the analysis of variance. MRA compares two parts of variance, firstly how much of the total variance that depends on the independent variables (the regression). This is compared with the part of the variance that remains unexplained by the independent variables (the residual). To have a higher probability of reaching statistical significance, a larger part of the variance should be in the regression and a smaller part in the residual (Djurfeldt et al., 2018).

In order to perform an MRA, all items were categorized into one of the independent variables. Based on these categorizations, indices were created, one index for each independent variable. However, before the indices were created and any further analysis was performed, the internal consistency of each variable was measured by examining the Cronbach's alpha value. When creating an index, one wants the variables to covary to a fairly high extent, as this means that they measure the same theoretical phenomenon. It is therefore important to test the reliability (Barmark & Djurfeldt, 2020). For Cronbach's alpha, it is desirable that the value is above 0.7 (Barmark & Djurfeldt, 2020; Pallant, 2020). However, Pallant (2020) explains that Cronbach's alpha is a sensitive value, meaning that it can be affected by the number of items. A low number of items can lead to a low Cronbach's alpha value. If this is the case, it is appropriate to examine an additional reliability value, the Mean Inter-Item Correlation. Since this is a small study with a low number of items, it was considered relevant to also check this value. For a Mean Inter-Item Correlation value to be acceptable, it should be between 0.2 and 0.4. If neither the value of Cronbach's alpha nor Mean Inter-Item Correlation are within the acceptable range, further action is required (Pallant, 2020). If that is the case, Pallant (2020) suggests that one or more items can be deleted to improve the value. Hence, this was done if required. Thereafter, the descriptive statistics for each variable was analyzed to examine the mean and standard deviation scores. This provided a summary and overview of the data and revealed which variable had the highest average score on the Likert scale.

Before carrying out the MRA, the Pearson correlation matrix was investigated to determine the strength of the correlations between each of the independent variables and the willingness to receive personalized communication based on personal information. Pallant (2020) describes that an r-value above 0.5 indicates a strong correlation and an r-value below 0.29 indicates a weak correlation. If the value is outside these limits, Pallant (2020) suggests performing a multicollinearity test. In order to perform MRA, the variables should not be too highly correlated, as this may indicate multicollinearity. If the correlation is too high, it is difficult to determine the unique effect of each index on the dependent variable (Pallant, 2020). The values examined in a multicollinearity test is tolerance and variance inflation factor (VIF). Tolerance is a measure of how much of the variation in an independent variable cannot be explained by the other independent variables. If the value of tolerance is low (less than 0.10), it indicates that there may be problems with multicollinearity. The VIF value is related to tolerance, and a high VIF value (above 10) indicates possible multicollinearity (Pallant, 2020).

Once the above aspects were ensured, the Multiple Regression Analysis (MRA) could be conducted. The MRA determines which, if any of the independent variables, have a significant impact on the dependent variable. Furthermore, the analysis determines whether the proposed hypotheses are supported or not (Pallant, 2020; Djurfeldt et al., 2018). The tests and values that were included in the MRA were R Squared (R^2), ANOVA, Coefficients; beta coefficient value (β) and sig. value (p), and the Normal probability plot. The R^2 value explains how much of the variance in the dependent variable is explained by the model as a whole (Pallant, 2020). However, since our sample is relatively small, Pallant (2020) suggests that the Adjusted R^2 value serves a better estimation of the real population. Therefore, this is the value that we decided to refer to. In order to determine if the model and the Adjusted R^2 was statistically significant, the p-value presented in the ANOVA table was examined to ensure that it was below Pallant's (2020) recommendation of p<0.05.

The analysis continued by examining the Coefficient table to evaluate the beta coefficient value (β) and sig. value (p). The β-value explains how much impact each independent variable has on the dependent variable: a higher value indicates a higher impact regardless of whether it is a positive or negative number (Pallant, 2020). To evaluate which of the independent variables that had the most unique significant contribution to the dependent variable, the p-values in the Coefficient table were examined where a value p<0.05 is considered statistically significant (Djurfeldt et al., 2018). The last step of the MRA is the Normal Probability Plot, which is used to visualize the skewness of the distributed values. This is to ensure that the research model does not show any major deviations from normality. Preferably the dots should follow a diagonal line from bottom left corner to top right, as this means that the values are normally distributed without any significant deviations (Pallant, 2020). Lastly, when the MRA was conducted, the results were presented to test if our hypotheses are either supported or rejected. This was followed by a discussion of the results and analysis.

3.5 Methodological limitations

One methodological limitation of the study concerns the sampling method, which was based on recruiting participants through our personal social media accounts. Although this approach allowed convenient access to potential respondents, it made it challenging to obtain a result that is generalizable to a population larger than the study sample. Furthermore, as the survey was sent out through our own social media accounts, there is a risk that the respondents are quite like minded to each other and us. Hence, a negative consequence of using convenience sampling as a strategy is that it is most likely to contribute to a skewed distribution of the demographic disproportion among respondents, leading to ungeneralizable results. Thus, the sampling method creates problems in terms of reliability and validity of the study as another researcher probably will not be able to obtain the same results following our methodological steps (Bryman et al. 2012). Moreover, the sample size of the study was limited by time and financial resources, which combined with the sampling method, risks resulting in a small sample size. A small sample lacks the statistical power required to identify subtle effects or variations within the population. This reduces the reliability of the results, as larger sample sizes are preferable to ensure statistical robustness (Pallant, 2020).

Lastly, there is a limitation regarding the constructed research model of the study. Since the variables perceived risk and perceived benefit are included in the Privacy Calculus Model (PCM) and it is a model that has been used in several previous studies, it is most likely that these variables are more elaborate than our self-constructed variables. Hence, it is more likely that the PCM part of our model will have a greater impact on the dependent variable due to its prior evidence. To increase the validity of the study, an operationalization was conducted in order to translate our theoretical concepts and previous research into measurable constructs (see appendix 7.1). However, as our self-constructed variables have not been tested in this context before, nor have they been evaluated or revised, they may not have a statistically significant contribution to our dependent variable. If the translated concepts in the operationalization do not measure what they intend to measure, this can lead to incorrect results and create problems in terms of validity of the study (Djurfeldt et al., 2018).

3.6 Ethical considerations

Bryman (2012) emphasizes the importance of providing respondents of a survey with sufficient information about the study before they decide to participate. Therefore, the decision was made to include an information text for the questionnaire, outlining the study’s purpose and content. This text also assured respondents that they were completely anonymous and that their participation was voluntary. According to Bryman (2012), the anonymity of respondents is an important ethical measure to exclude the possibility of tracing an answer to a specific person, protecting their privacy. Moreover, anonymity enhances the chance that respondents will answer questions honestly. Finally, no sensitive topic was examined and it can therefore be assumed that none of the respondents were triggered or harmed by participating in the study.

4. Results and analysis

This section presents the study's obtained results from the conducted analyses in SPSS. Firstly, the demographic distribution of the respondents is presented, followed by a reliability analysis and descriptive statistics. Furthermore, correlation analyses and the conducted multiple regression analysis are presented. Lastly, a test of our posed hypotheses is reported.

4.1 Demographic analysis

Demographic characteristics					
Sample		**1: Netflix**		**2: Spotify**	
		Frequency	Percentage	Frequency	Percentage
Gender	Male	27	28.4	33	34.7
	Female	68	71.6	62	65.3
Age	20	2	2.1	2	2.1
	21	6	6.3	4	4.2
	22	5	5.3	8	8.4
	23	6	6.3	6	6.3
	24	13	13.7	10	10.5
	25	33	34.7	28	29.5
	26	10	10.5	13	13.7
	27	11	11.6	7	7.4
	28	3	3.2	6	6.3
	29	1	1.1	7	7.4
	30	5	5.3	4	4.2
Education*	Primary school	0	0	2	2.1
	High School	53	55.8	45	47.4
	University/College	39	41.1	44	46.3
	Higher Vocational Education	3	3.2	4	4.2
	No completed education	0	0	0	0
Occupation	Student	65	68.4	45	47.4
	Employed	29	30.5	47	49.5
	Unemployed	1	1.1	3	3.2
Residence	Sweden	95	100	95	100
	Other	0	0	0	0

*Referring to the highest level of finished education.

***Table 1**: Presenting the demographic characteristics of our sample.*

Since this study is an experiment with two separate surveys, and the participants were selected through a convenience sample, the demographic statistics differ slightly between the

Netflix and Spotify respondents. The survey contained demographic filter questions to ensure responses from only our population. However, as the survey did not include a filter question regarding if the respondent were a user of Netflix or Spotify, one outlier was identified in that aspect. One respondent to the Netflix survey had not used the service in the last year, so in order to receive a more distinct result, we eliminated that outlier in the dataset. This left us with a sample of 95 participants per survey.

The demographic statistics show that out of the 95 participants in the Netflix survey, 71.6 percent were female and 28.4 percent were male. In the Spotify survey, which also had 95 participants, there is a similar gender distribution, with 65.3 percent female and 34.7 percent male. This makes the female participants a majority in both cases. Furthermore, 25-year-olds were the most represented in both groups, accounting for 34.7 percent in the case of Netflix and 29.9 percent in the case of Spotify. The remaining age groups are relatively normally distributed around this median. The highest level of education completed by the respondents is mainly High School or University/College. Out of the 95 Netflix participants, 55.8 percent have completed High School and 41.1 percent have completed University/College. Out of the 95 Spotify participants, 47.4 percent have completed High School and 46.3 percent have completed University/College. Regarding the occupational demographics, the majority of the Netflix participants are students (68.4 percent) while the “employed” option had the most responses in the Spotify case (49.5 percent). However, the number of students in the Spotify survey was also a large proportion (47.7 percent). Regarding the participants' residency, the study's sample was limited to people living in Sweden. Therefore, 100 percent of the respondents in both cases were from Sweden.

With these statistics, one can establish that the average respondent in our sample in the case of Netflix is a female student at the age of 25 who has completed High School, presumably studying at a more advanced level. In the case of Spotify, the average respondent in our sample is a female at the age of 25 who has finished her High School or University/College studies and is now employed. As previously reflected in the methodology section, the sample demographic distribution is not representative to the true population due to several limitations.

4.2 Reliability Analysis

4.2.1 Internal Consistency

Constructs	**Sample 1: Netflix**				**Sample 2: Spotify**		
	Cronbach's alpha	Mean Inter-item Correlation	No. of items		Cronbach's alpha	Mean Inter-item Correlation	No. of items
WR	0.671	0.253	6	**WR**	0.581	0.206	6
PB	0.934	0.780	4	**PB**	0.929	0.767	4
PR	0.730	0.309	6	**PR**	0.720	0.298	6
TU	0.451	0.294	2	**TU***	0.307	0.182	2 (0)
TY*	0.239	0.080	3 (0)	**TY***	0.228	0.078	3 (0)
CO	0.665	0.286	5	**CO**	0.546	0.227	5
				TT**	0.191 (0.557)	0.051 (0.304)	5 (3)
		Total items:	**26 (23)**			**Total items:**	**31 (24)**

***** Removed due to unacceptable value

****** New variable combining trust and transparency

Table 2: *Reliability testing: Cronbach's Alpha and Mean Inter-item Correlation.*

The Cronbach's alpha value for the first two independent variables, perceived benefits (PB, Netflix: 0.934; Spotify: 0.929) and perceived risks (PB, N: 0.730; S: 0.720), was higher than 0.7 for both groups. Thus, they were both above the limit described by Barmark and Djurfeldt (2020) and Pallant (2020) as acceptable for Cronbach's alpha, and therefore no further reliability assessment was needed for these variables. The dependent variable willingness to receive personalized communication based on personal information (WR) did not reach an acceptable value for any of the groups. However, the Mean Inter-Item Correlation value was 0.253 for Netflix and 0.206 for Spotify, which is within the acceptable limit (0.2-0.4) according to Pallant (2020), meaning that it did not need further assessments either.

Neither the third variable trust (TU, N: 0.451; S: 0.307) nor the fourth variable transparency (TY, N: 0.239; S: 0.228) had an acceptable Cronbach's alpha value. However, this was quite expected as Pallant (2020) describes that Cronbach's Alpha is a sensitive value and risks being low if there are a low number of items, which is the case in this study. The transparency variable also failed to meet the acceptable Mean Inter-Item Correlation value for both groups (N: 0.080; S: 0.078). For trust, the Mean Inter-Item Correlation value for Spotify was 0.182, which is not within the acceptable range. However, the trust Mean Inter-Item Correlation value for Netflix was 0.294, which meant that it was accepted.

Since Spotify did not have acceptable values for trust nor transparency, further reliability evaluation needed to be done for those variables. As both trust and transparency had few items each, no item could be removed from either of them to achieve a better value. Therefore, the decision was made to merge trust and transparency into a new variable. This variable is referred to as TT (Trust+Transparency) in the table above and throughout the study. This variable was thus only used in the analysis of Spotify, since Netflix had an accepted value for trust. The decision to merge trust and transparency can be further supported by previous research from Karwatzki et al. (2017) and Cloarec et al. (2024a) who describes that transparency is an important factor in achieving trust. It can therefore be argued that they are associated with each other. However, despite the merge, the value for TT was not at an acceptable level for Cronbach's alpha (0.191), nor Mean Inter-Item Correlation (0.051). Therefore, two items were removed, TY1 and TY2 (see appendix 7.1), to achieve an acceptable value. By removing these items, Spotify achieved a Mean Inter-Item Correlation value of 0.304, which is an acceptable value.

The last independent variable context (CO) had a Cronbach's alpha value of 0.665 for Netflix and 0.546 for Spotify, which is too low. However, both groups had an acceptable value for the Mean Inter-Item Correlation (N: 0.286; S: 0.227), which means that no further reliability assessment was needed.

For Netflix, it was thus only transparency that did not have an acceptable Cronbach's alpha or Mean Inter-Item Correlation value. Removing an item did not increase the value either. Hence, there were no more options to increase the value. Therefore, the decision was made to remove the variable transparency from Netflix and not include it in further analyses. In Spotify, it was also removed in its individual form, but as described above, it was merged with trust.

4.3 Descriptive statistics

Sample 1: Netflix	**WR**	**PB**	**PR**	**TU**	**CO**
Mean	3.9508	3.8000	3.5087	4.0158	2.6547
Std. Deviation	0.6306	0.8973	0.7298	0.8361	0.6875
Sample 2: Spotify	**WR**	**PB**	**PR**	**TT**	**CO**
Mean	3.9641	3.7210	3.6315	4.1579	3.0252
Std. Deviation	0.5729	1.0011	0.7535	0.6447	0.5451

Table 3: *Mean and standard deviation of the study index variables.*

The table above presents descriptive statistics for the mean and standard deviation of the study's constructed index variables, and provides an overview of the respondents' answers. The response options were designed according to a five-point Likert scale where 1 referred to “Strongly Disagree” and 5 referred to “Strongly Agree”. The table shows that TU (4.0158) for Netflix and TT (4.1579) for Spotify had the highest mean score, and were thus closest to “Strongly Agree”. However, they were both closer to 4 which stands for “Agree”. CO had the lowest mean score for both groups (N: 2.6527; S: 3.0252) and the values were thus closest to response option 3 meaning “Neither agree nor disagree”. Furthermore, CO was, together with WR, the indices with the lowest standard deviation for both groups, which means that these variables had the lowest dispersion of responses. Overall, the mean for Netflix was between 2.6527-4.0158, and for Spotify it was between 3.0252-4.1579. In terms of standard deviation, for Netflix it ranged between 0.6306-0.8973, and for Spotify it ranged between 0.5451-1.0011.

4.4 Correlation

4.4.1 Pearson Correlation

	Sample 1: Netflix						Sample 2: Spotify				
Variable	**WR**	**PB**	**PR**	**TU**	**CO**	Variable	**WR**	**PB**	**PR**	**TT**	**CO**
WR	1	0.642	0.238	0.787	0.248	**WR**	1	0.747	-0.018	0.855	0.622
PB	0.642	1	-0.146	0.430	0.515	**PB**	0.747	1	-0.405	0.629	0.744
PR	0.238	-0.146	1	-0.005	-0.514	**PR**	-0.018	-0.405	1	-0.190	-0.321
TU	0.787	0.430	-0.005	1	0.350	**TT**	0.855	0.629	-0.190	1	0.568
CO	0.248	0.515	-0.514	0.350	1	**CO**	0.622	0.744	-0.321	0.568	1

Table 4: *Pearson Correlation values between all constructs in the model. The most important values are highlighted in grey.*

When examining the correlation between the dependent (WR) and independent (PB, PR, TT & CO) variables in the Pearson Correlation table, the results showed that one of the independent variables in the case of Spotify (PR), and two of the independent variables in the case of Netflix (PR and CO) had an *r*-value below 0.29. This indicates that they correlate weakly with the dependent variable (WR). The remaining independent variables in both cases (TT, CO & PB) had an *r*-value above 0.5, meaning that they had a strong correlation with the dependent variable (WR). However, it was found that the independent variables PB and CO in the case of Spotify had an *r*-value above 0.7, which could indicate that there is a multicollinearity between them. Thus, a decision was made to conduct a multicollinearity test for Spotify to investigate it further (Pallant, 2020). The multicollinearity test is presented below.

4.4.2 Multicollinearity Test

Model	**Sample 2: Spotify**	
	Tolerance	**VIF**
PB	0.354	2.828
PR	0.827	1.209
TT	0.577	1.733
CO	0.428	2.334

Table 5: *Multicollinearity test for Spotify with Tolerance and VIF value.*

The multicollinearity test showed acceptable results, as seen in the table above. The tolerance values were significantly higher than 0.10, and the VIF values were not close to being above 10 for each independent variable. This implies that our model is a good fit for conducting a multiple regression analysis (Pallant, 2020).

4.5 Multiple regression analysis (MRA)

4.5.1 R Squared and ANOVA

Model	**Sample 1: Netflix**			**Sample 2: Spotify**		
1	**R**	**R Squared**	**Adjusted R Square**	**R**	**R Squared**	**Adjusted R Square**
	0.907	0.823	0.815	0.936	0.876	0.870

Table 6: *R Squared and Adjusted R Square results.*

The analysis began by evaluating how much of the variance in the dependent variable is explained by the model, by examining the R Squared value (R^2). As seen in the table above, the R^2 value is 0.823 for Netflix and 0.876 for Spotify, meaning that the model explains 82.3 percent respectively 87.6 percent of the variance in WR. However, worth noting is that the study's sample is relatively small (95 valid responses per survey), which indicates that the R^2 value tends to be optimistic. Therefore, the Adjusted R^2 corrects this value to provide a better estimate of the population (Pallant, 2020). Thus, the reported value will be the Adjusted R^2

which was 0.815 for Netflix and 0.870 for Spotify, meaning that the independent variables (PB, PR, TT and CO) explain 81.5 percent (N) respectively 87.0 percent (S) of the variance in the dependent variable (WR). Even the Adjusted R^2 value can be considered unreasonably high, which may indicate that the model overfits our specific data set but does not generalize well to new data, creating problems in terms of validity of the study (Pallant, 2020). However, as this is the obtained results, these values will be used with awareness of their limitations. The *p*-value (Sig.) shown in the ANOVA table (table 7) determines whether the model is statistically significant, which is the case if the value is $p < 0.05$. As our ANOVA table shows a value of *Sig.* < 0.001 (which is equal to $p < 0.05$) in both groups, our model is statistically significant (Pallant, 2020).

ANOVA

Model		**Sample 1: Netflix**		**Sample 2: Spotify**	
		Sum of Squares	**Sig.**	**Sum of Squares**	**Sig.**
1	**Regression**	1107.996	0.001	972.618	0.001
	Residual	237.752		138.287	
	Total	1345.747		1110.905	

Table 7: *Analysis of Variance - ANOVA.*

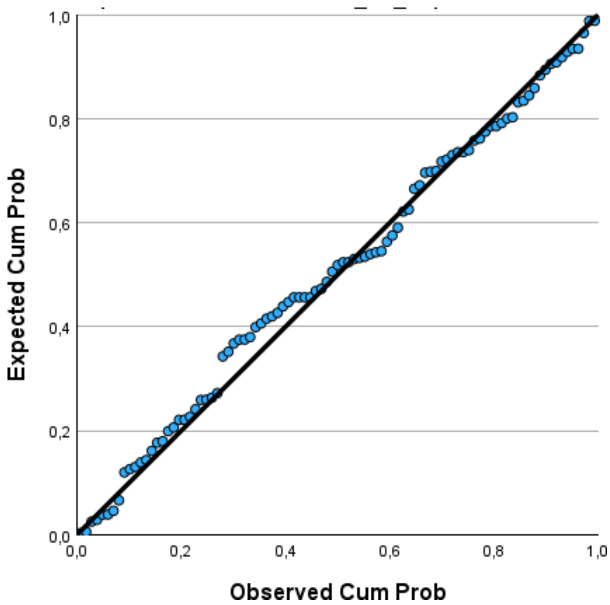
4.5.2 Coefficients

	Sample 1: Netflix			Sample 2: Spotify	
	Standardized Coefficients	**Sig.**		**Standardized Coefficients**	**Sig.**
Constructs	Beta	p <0.05	**Constructs**	Beta	p <0.05
PB	0.446	0.001	**PB**	0.469	0.001
PR	0.278	0.001	**PR**	0.295	0.001
TU	0.615	0.001	**TT**	0.602	0.001
CO	-0.055	0.378	**CO**	0.026	0.647

Table 8: *Presenting the coefficient values: β-value and p-value.*

As the model above shows, TT had the most unique impact on the willingness to receive personalized communication based on personal information in the case of Spotify, with a beta value of β=0.602 and a significance value of 0.001. In the case of Netflix, TU was the variable with the most unique impact, showing a beta value of β=0.615 and a significance value of 0.001. The variables PB and PR also had an acceptable significance value in both cases (N: 0.001; S: 0.001), meaning that these two also have a unique, but not as strong, impact on the dependent variable. However, CO for both groups had low beta values and did not show a p<0.05, indicating that it does not have a unique or significant contribution to the dependent variable (Pallant, 2020).

4.5.3 Normal Probability Plots



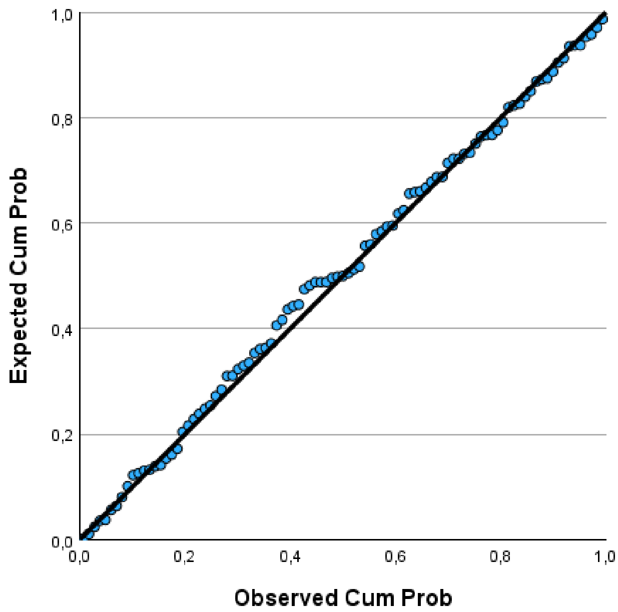


Figure 2: *Netflix Normal Probability Plot.* ***Figure 3:*** *Spotify Normal Probability Plot.*

Neither of the models above shows any major deviation from normality as the dots follow the diagonal line relatively closely. There is some deviation for Netflix's graph between 0.3 and 0.45, but it is not enough to cause issues for the study.

4.6 Hypothesis testing

		Sample 1: Netflix			Sample 2: Spotify		
Hypotheses	**Effect**	**Beta**	**Sig.**	**Result**	**Beta**	**Sig.**	**Result**
H1	PB → WR	0.446	0.001	Supported	0.459	0.001	Supported
H2	PR → WR	0.278	0.001	Supported	0.295	0.001	Supported
H3	TU → WR	0.615	0.001	Supported	-	-	Rejected
H4	TY → WR	-	-	Rejected	-	-	Rejected
H5	CO → WR	-0.055	0.378	Rejected	0.026	0.647	Rejected
H6*	TT → WR	-	-	N/A	0.602	0.001	Supported

*New hypothesis testing the variable TT

Table 9: *Table presenting the supported and rejected hypotheses based on the β-values and p-values.*

The table above shows the expected effect of each hypothesis, the standardized beta coefficient, the *p*-value and the decision to support or reject each hypothesis. The results show that for Netflix, three hypotheses were supported (H1, H2 & H3) and two hypotheses were rejected (H4 & H5). Since a new variable was created for Spotify (TT), a new hypothesis was also added (H6) to test it. The results for Spotify show that three hypotheses were supported (H1, H2 & H6), while three were rejected (H3, H4 & H5).

H1 and H2 had a significance value of 0.001 for both Netflix and Spotify, thus in line with the accepted value of $p<0.05$, meaning that they have a statistically significant impact on the dependent variable WR. Hence, H1 and H2 were supported for both Netflix and Spotify. The beta value for PB was $\beta=0.446$ for Netflix and $\beta=0.459$ for Spotify, which was the second highest beta value for both and thus the variable with the second strongest unique contribution on WR.

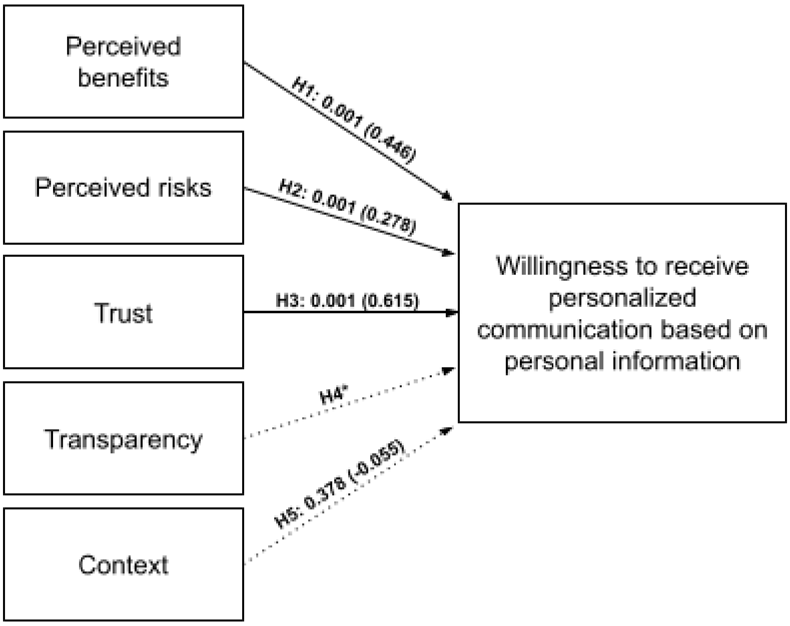
Since the variable TU was not tested on its own for Spotify, but was added to a new variable (TT), H3 could not be tested for Spotify, and was therefore rejected. However, for Netflix it was possible to test the TU variable, and the table shows that H3 had a *p*-value of 0.001, indicating that H3 has a significant impact on WR. The beta value for H3 was $\beta=0.615$, which was the highest beta value for all of Netflix's variables. This means that TU has the greatest unique impact on WR when it comes to Netflix, further strengthening the significance of H3.

Given that the independent variable TY in H4 did not show acceptable values in the reliability test (Table 2), we decided as previously mentioned, not to proceed with that variable on its own in the analysis. As a result, H4 could not be tested for either Netflix nor Spotify and was therefore rejected. Furthermore, H5 was also rejected in both cases since the significance value was above the accepted value of $p<0.05$. However, it is worth noting that the significance value for Netflix in H5 was closer to $p<0.05$ than it was for Spotify, but it is still not sufficient to establish that the result is due to our observations and not to chance. Moreover, both Netflix and Spotify had low beta values in H5 (N: $\beta=-0.055$; S: $\beta=0.026$), meaning that unique contribution to WR would still be low even if it was significant.

As mentioned in section 4.2.1, we decided to combine the independent variables TY and TU in the case of Spotify in order to achieve an acceptable Mean Inter-Item Correlation value. Therefore, the new variable TT needed to be created. As a result, an additional H6 hypothesis was created. H6 assumes that trust in combination with transparency will have a significant positive impact on our dependent variable WR. When examining the beta value and p-value of H6, one can establish that this hypothesis has the strongest unique contribution on the dependent variable in the case of Spotify, with values of β=0.602 and p=0.001.

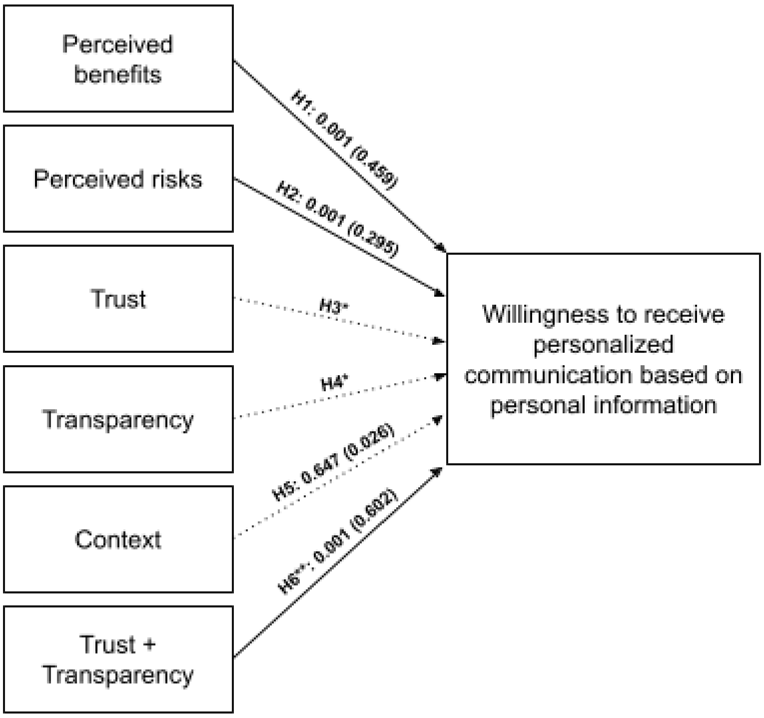
4.6.1 Research Models Results

Presented below is the study's research model with the tested hypotheses. The rejected hypotheses have a dotted line.



*Hypothesis not tested due to unacceptable internal reliability

Figure 4: *Netflix research model with hypotheses results: P-values with Standardized Beta Coefficients inside brackets.*



*Hypothesis not tested due to unacceptable internal reliability

**Additional hypothesis for the new variable TT

Figure 5: *Spotify research model with hypotheses results: P-values with Standardized Beta Coefficients inside brackets.*

5. Discussion and conclusion

In this chapter, the results from the SPSS analyzes are discussed in relation to previous research of the study. This is followed by a conclusion where the research question is answered. Lastly, our suggestions for future research are presented.

5.1 Discussion

The aim of this thesis was to examine which factors influence the willingness to receive personalized communication on streaming services. Furthermore, the study aimed to investigate the intention to disclose personal information, as it is a prerequisite for personalized communication. This was accomplished by employing the Privacy Calculus Model (PCM). In addition to the initial factors of perceived benefit and perceived risk, three supplementary factors were integrated: trust, transparency, and context. The reason for this was that according to previous research, these factors were also assumed to have an impact on the dependent variable.

The results of the MRA presented in table 9 showed that the independent variables perceived benefits (PB) and perceived risks (PR) both had a significant influence on the explanation of our sample's willingness to receive personalized communication based on personal information. This applied to both Netflix and Spotify. In other words, our sample considers that perceived benefits of personalized communication lead to an increase in their willingness to receive it. At the same time, they believe that the perceived risks of personalized communication lead to a decreased willingness to receive it. This contradiction could be explained by the Personalization Privacy Paradox (PPP), which implies that people desire a personalized experience due to the benefits that come with it, while at the same time being concerned about their personal information and privacy online (Wang et al., 2024).

Cloarec et al. (2024b) explain that individuals evaluate the costs and benefits of sharing personal information for personalized recommendations, facing a trade-off. Disclosing sensitive data can lead to highly personalized offers, and in our case, enhance ease of use, content satisfaction and time savings when using a streaming service. However, opting not to share such information reduces these benefits and results in more standardized content. Thus, our results are consistent with Hayes et al. (2021) study on the PCM, which describes that consumers make a trade-off between benefits and risks before deciding to share personal information. We expected the PCM to be consistent with our findings as it is a well-developed model that has been tested in several other studies. However, we anticipated that there would be a greater difference between the groups, which turned out to be true only to a low degree. The respondents exposed to the Spotify condition valued the benefits slightly higher, but the difference was not large enough to be meaningful. This is probably due to our low number of respondents in each survey, which makes it difficult to confirm the accuracy of the results.

Furthermore, in the Netflix case, it was found that the independent variable trust (TU) had the strongest unique contribution to the dependent variable. Hence, our sample values trust the most in their decision to disclose personal information to receive personalized communication. This indicates that a high level of trust outweighs the risks of disclosing personal information, and thus makes one more willing to receive personalized communication. Bleier and Eisenbeiss (2015) highlight that trust plays a vital role in individuals' perceptions of personalization. Similar to our findings, they argue that higher levels of trust in a company are associated with lower levels of privacy concerns about hidden agendas and mishandling of personal information. As one of our two items regarding trust were “Overall, I have more trust in Swedish companies than in foreign companies”, one could argue that Bhattacharya et al. (2023) findings regarding country of origin (COO) and home country bias are applicable. Bhattacharya et al. (2023) describe that people tend to have a higher level of trust in companies established in their home country, compared to foreign companies, and that people have more trust in domestic companies to process their personal data correctly. Therefore, our results suggest that our sample has a high level of trust in Swedish companies, and thus are more cautious about sharing their personal information with Netflix, which is a US company.

To compare these results with the Spotify case, it is important to remember that the variables trust and transparency had to be combined to ensure acceptable internal reliability. As mentioned in section 2.4.4, Karwatzki et al. (2017) and Cloarec et al. (2024a) explain that transparency is often seen as a crucial factor in building trust. Therefore, our decision to combine trust with transparency to create the new independent variable TT was considered justifiable. When investigating the values of TT in the case of Spotify, similar values to TU for Netflix were found. Thus, trust merged with transparency was the independent variable in the case of Spotify which had the strongest unique contribution to the dependent variable. The reason why it was not possible to form an index for trust alone may be due to various factors. One reason could be that people do not reflect as much on the risks of disclosing personal information to a Swedish company. Similarly, people may not reflect as much on the role of trust because the disclosure is more unconscious. However, as transparency could be included in the index, it may indicate that the respondents perceive more trust in Spotify as they perceive them as more transparent in their communication about data usage. This is in line with Karwatzki et al. (2017) and Cloarec et al. (2024a) research, that clear and user-friendly communication about how personal data is processed tends to increase trust in the company. Worth noting is that the items for both trust and transparency can be considered too few to draw any accurate conclusions, creating problems in terms of the validity of our study.

The results imply that our sample's willingness to receive personalized communication based on personal information is mainly influenced by the perceived benefits, perceived risks and trust (Netflix), or trust in combination with transparency (Spotify). Hence, the context variable showed little to no explanation of the dependent variable making it insignificant for the willingness to receive personalized communication based on personal information. However, as the context permeated each survey implicitly due to its experimental design, it can be argued that it may still have had an impact, even if the values in the hypothesis testing do not confirm it. When examining the results from the MRA and the mean score for the context variable, the values are relatively similar for both Netflix and Spotify. One possible explanation for this could be that both Netflix and Spotify are two well-known companies. Thus, they have an inherent trust among their users, and therefore users are more likely to disclose their personal information to them. This explanation is consistent with Augirre et al. (2016) study, which emphasizes that people are more comfortable with sharing their personal data to a well-known company. Thus, it may be that the context would have had a greater impact if we had compared a well-known and a less well-known company instead. Presumably, a less well-known company has not managed to build as much trust with its consumers. Thus, our results would probably be different if that comparison had been made.

In retrospect, the items in the context variable should have been more carefully considered to truly measure the impact of context on the dependent variable. Furthermore, the rejected hypotheses could be due to limited previous research within this field regarding streaming services. Hence, this hypothesis does not have as solid foundation as the remaining variables of the study. Another explanation to the rejected hypotheses is emphasized by the mean scores of the context variable. In both groups, the mean scores were close to 3, meaning that the respondents in average answered “Neither agree nor disagree”. This could either mean that they had no opinion about the context, or that they did not understand our questions. In any case, it makes the context variable insignificant to explain the variation in the dependent variable.

5.2 Conclusion

To conclude, this study aimed to investigate which of the factors perceived benefits, perceived risks, trust, transparency and context that had the most significant impact on the willingness to receive personalized communication based on personal information. Furthermore, the aim was to examine whether there was any difference in the willingness depending on the streaming service. The study's formulated research question was:

> *Which factor of perceived benefits, perceived risks, trust, transparency and context have the most impact on streaming service users' willingness to receive personalized communication based on personal information?*

To answer this and fulfill the aim of the study, five hypotheses were formulated based on previous literature.

When conducting the analysis in SPSS, some of the variables had to be merged to reach an acceptable internal consistency, resulting in a sixth hypothesis. However, the sixth hypothesis was only applicable for Spotify. The analysis showed that three hypotheses per sample were statistically supported. In the case of Netflix, H1, H2 and H3 were supported. In the case of Spotify, H1, H2 and H6 were supported.

To answer the research question regarding which factor had the most impact on users' willingness to receive personalized communication, trust was the independent variable that stood out statistically for Netflix. For Spotify, the merged variable of trust and transparency had the most impact. Whether there is any difference depending on the context is difficult to answer. On one hand, H4 was rejected, meaning that it is not statistically significant to determine its impact. On the other hand, as the context permeated the entire study given its experimental design and given that the results for Netflix and Spotify were slightly different, there is a possibility that the context may have had an influence. However, this could not be confirmed statistically through our study. Therefore, the only conclusion can be that context did not influence users' willingness to receive personalized communication based on personal information.

Our study provides valuable insights for communicators and companies using personalization strategies in their digital marketing communications. The study suggests that companies should prioritize strategies aimed at building trust to encourage customers to be willing to share their personal data. Furthermore, it is important that companies take responsibility and are clear about how they process their customers' personal data. Additionally, reducing customers' perceived risks is crucial in fostering their willingness to share personal information, enhancing their perceived benefits. Today's consumers desire and expect communication from companies to be tailored to them. Moreover, there is evidence that these strategies increase profits for businesses. Hence, given that personalization strategies require personal data to be successful, addressing these issues should be a top priority for companies to meet customer needs and stay competitive in the emerging era of digitalization.

5.3 Suggestions to future research

To obtain more generalizable results, a suggestion for future research is to conduct a larger-scale quantitative study. Although our study provided valuable insights, its scope was limited due to resource constraints, resulting in a small sample size. By expanding the sample size, future researchers can increase the representativeness of the results and improve the generalisability of the outcomes to a broader population. Furthermore, it would be interesting to investigate other demographic groups, such as older people, as this could provide valuable insights into possible age-related differences in willingness to receive personalized communication. Moreover, as our study predominantly included women, it would be interesting to examine a more diverse gender distribution to determine if there is any difference based on gender. Furthermore, examining the opinions of people other than Swedes may contribute to broader insights, as cultural differences affect how privacy is valued and prioritized.

A further suggestion for future research is to adopt a qualitative approach, for example through interviews or focus groups. This would allow for a more nuanced understanding of the underlying motivations behind people's attitudes towards receiving personalized communication. As privacy is a complex and ambiguous subject, the qualitative approach could contribute to understanding the phenomenon from different perspectives and in more depth.

Lastly, future research could benefit from further exploring the context. As previous research highlights that well-known companies tend to have an inherent trust among their stakeholders, comparing a well-known company with a lesser known company could shed light on how different levels of trust impacts people's perceptions. Such a comparison could go beyond trust to include other contextual factors, as for example industry norms and brand reputation.

6. Bibliography

Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., & Wetzels, M. (2015). Unraveling the Personalization Paradox: The Effect of Information Collection and Trust-Building Strategies on Online Advertisement Effectiveness. *Journal of Retailing*, *91*(1), 34–49. https://doi.org/10.1016/j.jretai.2014.09.005

Aguirre, E., Roggeveen, A.L., Grewal, D., & Wetzels, M. (2016). The personalization-privacy paradox: implications for new media. *Journal of Consumer Marketing*, *33*(2), 98-110. https://doi.org/10.1108/JCM-06-2015-1458

Alimamy, S., & Gnoth, J. (2022). I want it my way! The effect of perceptions of personalization through augmented reality and online shopping on customer intentions to co-create value. *Computers in Human Behavior*, *128*(2022), 107105. https://doi.org/10.1016/j.chb.2021.107105

Ameen, N., Hosany, S., & Paul, J. (2022). The personalisation-privacy paradox: Consumer interaction with smart technologies and shopping mall loyalty. *Computers in Human Behavior*, *126*(2022), 106976. https://doi.org/10.1016/j.chb.2021.106976

Anthony, J. T., Christian, G. E., Evanlim, V., Lucky, H., & Suhartono, D. (2022). The Utilization of Content Based Filtering for Spotify Music Recommendation. *2022 International Conference on Informatics Electrical and Electronics (ICIEE), 2022, 1-4.* https://doi.org/10.1109/ICIEE55596.2022.10010097

Barmark, M., & Djurfeldt, G. (2020). *Statistisk verktygslåda 0 : att förstå och förändra världen med siffror* (2nd ed.). Lund: Studentlitteratur.

Behera, G., & Nain, N. (2023). Collaborative Filtering with Temporal Features for Movie Recommendation System. *Procedia Computer Science*, *218*(2023), 1366-1373. https://doi.org/10.1016/j.procs.2023.01.115

Behera, R.K., Gunasekaran, A., Gupta, S., Kamboj, S., & Bala, P.K. (2020). Personalized digital marketing recommender engine. *Journal of Retailing and Consumer Services, 53* (2020), 101799. https://doi.org/10.1016/j.jretconser.2019.03.026

Bender, M., Gal-Or, E., & Geylani, T. (2022). The emergence of streaming and its impact on pricing and product strategies of content providers. *Managerial & Decision Economics, 43*(6), 2148-2171. https://doi.org/10.1002/mde.3515

Betzing, J.H., Tietz, M., Brocke, J., & Becker, J. (2020). The impact of transparency on mobile privacy decision making. *Electron Markets, 30*(3), 607–625. https://doi.org/10.1007/s12525-019-00332-3

Bhattacharya, S., Sharma, R. P., & Gupta, A. (2023). Does e-retailer's country of origin influence consumer privacy, trust and purchase intention?. *Journal of Consumer Marketing, 40*(2), 248–259. https://doi.org/10.1108/JCM-04-2021-4611

Bile Hassan, I., Murad, M.A.A., El-Shekeil, I., Liu, J. (2022). Extending the UTAUT2 Model with a Privacy Calculus Model to Enhance the Adoption of a Health Information Application in Malaysia. *Informatics, 9*(2), 31. https://doi.org/10.3390/informatics9020031

Bleier, A., & Eisenbeiss, M. (2015). The Importance of Trust for Personalized Online Advertising. *Journal of Retailing, 91*(3), 390-409. https://doi.org/10.1016/j.jretai.2015.04.001

Bobadilla, J., & Gutiérrez, A. (2024). Wasserstein GAN-based architecture to generate collaborative filtering synthetic datasets. *Applied Intelligence: The International Journal of Research on Intelligent Systems for Real Life Complex Problems*, 54(3), 2472–2490. https://doi.org/10.1007/s10489-024-05313-4

Bol, N., Kruikemeier, S., Boerman, S.C., Strycharz, J., de Vreese, C.H., Dienlin, T., Sax, M., & Helberger, N. (2018). Understanding the effects of personalization as a privacy calculus: Analyzing self-disclosure across health, news, and commerce contexts. *Journal of Computer-Mediated Communication, 23*(6), 370–388. https://doi.org/10.1093/jcmc/zmy020

Boyle, M., & Schmierbach, M. (2020). *Applied Communication Research Methods: Getting Started as a Researcher* (2nd ed.). Routledge.

Boyle, M., & Schmierbach, M. (2024). *Applied Communication Research Methods: Getting Started as a Researcher* (3rd ed.). New York: Routledge. https://doi.org/10.4324/9781003316831

Bryman, A. (2012). *Social research methods* (4 ed.). Oxford University Press.

Canhoto, A.I., Keegan, B.J., & Ryzhikh, M. (2023). Snakes and Ladders: Unpacking the Personalisation-Privacy Paradox in the Context of AI-Enabled Personalisation in the Physical Retail Environment. *Information Systems Frontiers: A Journal of Research and Innovation,* (2023), 1-20. https://doi.org/10.1007/s10796-023-10369-7

Chen, H.T. (2018). Revisiting the Privacy Paradox on Social Media With an Extended

Privacy Calculus Model: The Effect of Privacy Concerns, Privacy Self-Efficacy, and Social Capital on Privacy Management. *American Behavioral Scientist, 62*(10), 1392–1412. https://doi.org/10.1177/0002764218792691

Clark, T., Bryman, A., Foster, L., & Sloan, L. (2021). *Bryman's social research methods* (6th edition). Oxford University Press.

Cloarec, J., Cadieu, C., & Alrabie, N. (2024a). Tracking technologies in eHealth: Revisiting the personalization-privacy paradox through the transparency-control framework. *Technological Forecasting & Social Change*, *200*(2024), 123101 https://doi.org/10.1016/j.techfore.2023.123101

Cloarec, J., Meyer-Waarden, L., & Munzel, A. (2024b). Transformative privacy calculus: Conceptualizing the personalization-privacy paradox on social media. *Psychology & Marketing*, 1-23. https://doi.org/10.1002/mar.21998

Djurfeldt, G., Larsson, R., & Stjärnhagen, O. (2018). *Statistisk verktygslåda 1 : samhällsvetenskaplig orsaksanalys med kvantitativa metoder* (3rd ed.). Lund: Studentlitteratur.

Falkheimer, J., & Heide, M. (2018). *Strategic Communication - An Introduction*. Routledge

Fergusson, A. (2016). Designing online experiments using Google forms+ random redirect tool. Retrieved from https://teaching.statistics-is-awesome.org/designing-online-experiments-using-google-forms-random-redirect-tool

Fierro, I., Cardona Arbelaez, D.A., & Gavilanez, J. (2017). Digital marketing: a new tool for international education. *Pensamiento & Gestión*, (43), 240-260. https://doi.org/10.14482/pege.43.10594

Gironda, J. T., & Korgaonkar, P.K. (2018). iSpy? Tailored versus Invasive Ads and Consumers' Perceptions of Personalized Advertising. *Electronic Commerce Research and Applications, 29*(2018), 64-77. https://doi.org/10.1016/j.elerap.2018.03.007

Gong, Y., Schroeder, A., Pan., B., Sundar, S.S., & Mowen. A.J. (2024). Does algorithmic filtering lead to filter bubbles in online tourist information searches?. *Information Technology & Tourism, 26*(1), 183-217. https://doi.org/10.1007/s40558-023-00279-4

Hayes, J.L., Brinson, N.H., Bott, G.J., & Moeller, C.M. (2021). The Influence of Consumer–Brand Relationship on the Personalized Advertising Privacy Calculus in Social Media. *Journal of Interactive Marketing, 55*(2021), 16-30. https://doi.org/ 10.1016/j.intmar.2021.01.001

Hirschprung, R.S. (2023). Is the Privacy Paradox a Domain-Specific Phenomenon.

Computers, 12(8), 156. https://doi.org/10.3390/computers12080156

Internetstiftelsen. (2022). *Strömmande videotjänster fortsätter att öka bland internetanvändarna.* (Retrieved April 26, 2024). https://svenskarnaochinternet.se/rapporter/svenskarna-och-internet-2022/strommande-och-digitala-medietjanster/

Jiang, X., Goh, T.T., & Liu, M. (2022). On Students' Willingness to Use Online Learning: A Privacy Calculus Theory Approach. *Frontiers in Psychology, 13*(2022). https://doi.org/10.3389/fpsyg.2022.880261

Karwatzki, S., Dytynko, O., Trenz, M., & Veit, D. (2017). Beyond the Personalization-Privacy Paradox: Privacy Valuation, Transparency Features, and Service Personalization. *Journal of Management Information Systems*, 34(2), 369–400. https://doi.org/10.1080/07421222.2017.1334467

Kawaf, F., Montgomery, M., & Thuemmel, M. (2023). Unpacking the privacy-personalisation paradox in GDPR-2018 regulated environments: consumer vulnerability and the curse of personalisation. Information Technology and People, 37(4), 1674-1695. https://doi.org/10.1108/ITP-04-2022-0275

Khoo, O. (2023). Picturing Diversity: Netflix's Inclusion Strategy and the Netflix Recommender Algorithm (NRA). *Television & New Media*, *24*(3), 281-297. https://doi.org/10.1177/15274764221102864

Kronemann, B., Kizgin, H., Rana, N., & K. Dwivedi, Y. (2023). How AI encourages consumers to share their secrets? The role of anthropomorphism, personalisation, and privacy concerns and avenues for future research. *Spanish Journal of Marketing -ESIC, 27*(1), 3-19. https://doi.org/10.1108/SJME-10-2022-0213

Li, S., Peng, K., Zhu, B., Li, Z., Zhang, B., Chen, H., & Li, R. (2023). Research on Users' Privacy-Sharing Intentions in the Health Data Tracking System Providing Personalized Services and Public Services. *SUSTAINABILITY,* 15(22), 15709. https://doi.org/10.3390/su152215709

McKee, K.M., Dahl, A.J., & Peltier, J.W. (2023). Gen z's personalization paradoxes: A privacy calculus examination of digital personalization and brand behaviors. *Journal of Consumer Behaviour, 23*(2), 405-422. https://doi.org/10.1002/cb.2199

Mehta, R., Kulkarni, U. (2020). Impact of Personalized Social Media Advertisements on Consumer Purchase Intention. *Annals of Dunarea de Jos University. Fascicle I : Economics and Applied Informatics, 26*(2), 15-24. https://doi.org/10.35219/eai15840409101

Mohammed, A. A., & Rozsa, Z. (2024). Consumers' intentions to utilize smartphone diet applications: an integration of the privacy calculus model with self-efficacy, trust and experience. British Food Journal, 126(6), 2416-2437. https://doi.org/10.1108/BFJ-11-2023-0989

Netflix. (2022, November 1). *Sekretessförklaring.* (Retrieved April 10, 2024). https://help.netflix.com/legal/privacy

Netflix. (n.d.a). *The Story of Netflix*. (Retrieved April 10, 2024). https://about.netflix.com/en

Netflix. (n.d.b). *What is Netflix?.* (Retrieved April 10, 2024). https://help.netflix.com/en/node/412

Pallant, J. (2020). *SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). Open University Press, McGraw-Hill Education

Palmatier, R.W., Kumar, V., & Harmeling, C.M. (2018). *Customer Engagement Marketing* (1st ed.). Springer International Publishing.

Plangger, K., & Montecchi, M. (2020). Thinking Beyond Privacy Calculus: Investigating Reactions to Customer Surveillance. *Journal of Interactive Marketing, 50*(2020), 32-44. https://doi.org/10.1016/j.intmar.2019.10.004

Pukas, A.. (2022). Hyper-Personalization as a Customer Relationship Management Tool in a SMART Organization. *Problemy Zarządzania, 20*(3 (97)), 95–108. https://doi.org/10.7172/1644-9584.97.5

Sharma, K., & Aggarwal, S. (2023). *Digital Marketing Outreach : The Future of Marketing Practices* (1st ed.). Routledge.

Shim, H., & Yeon, J. (2022). Two-Facedness of Netflix Users?: Privacy Paradox with Privacy Insensitivity in Using Video Streaming Service. *SSRN*. Online pre-publication. http://dx.doi.org/10.2139/ssrn.4107152

Spotify. (n.d.). *About Spotify*. (Retrieved April 10, 2024). https://newsroom.spotify.com/company-info/

Spotify. (2023, December 5). *Spotifys integritetspolicy.* (Retrieved April 10, 2024). https://www.spotify.com/se/legal/privacy-policy/#3-personuppgifter-som-vi-samlar-in-om-dig

Tang, Y., & Ning, X. (2023). Understanding user misrepresentation behavior on social apps: The perspective of privacy calculus theory. *Decision Support Systems February, 165*(2023), 113881. https://doi.org/10.1016/j.dss.2022.113881

Tofalvy, T., & Koltai, J. (2023). “Splendid Isolation”: The reproduction of music industry inequalities in Spotify’s recommendation system. *New Media and Society*, 25(7), 1580-1604. https://doi.org/10.1177/14614448211022161

Trost, J., & Hultåker, O. (2016). *Enkätboken* (5th ed.). Lund: Studentlitteratur

Tsigkari, D., Iosifidis, G., & Spyropoulos, T. (2024). Quid Pro Quo in Streaming Services: Algorithms for Cooperative Recommendations. *IEEE Transactions on Mobile Computing, 23*(2), *1753–1768*. https://doi.org/10.1109/TMC.2023.3240006

Twilio Segment. (2023). *The State of Personalization report*. https://segment.com/state-of-personalization-report/

Wang, Y., Zhu, J., Liu, R., & Jiang, Y. (2024). Enhancing recommendation acceptance: Resolving the personalization–privacy paradox in recommender systems: A privacy calculus perspective. *International Journal of Information Management, 76*(2024), 102755. https://doi.org/10.1016/j.ijinfomgt.2024.102755

Wenemark, M. (2023). *Enkätmetodik - att planera och genomföra en undersökning* (1st ed.). Liber AB.

Wrench, J.S., Thomas-Maddox, C., Richmond V.P., & McCroskey, J.C. (2013). *Quantitative research methods for communication : a hands-on approach*. New York: Oxford University Press.

Yildiz, E., Güngör Şen, C., & Işık, E.E. (2023). A Hyper-Personalized Product Recommendation System Focused on Customer Segmentation: An Application in the Fashion Retail Industry. *Journal of Theoretical and Applied Electronic Commerce Research, 18*(1), 571-596. https://doi.org/10.3390/jtaer18010029

7. Appendices

7.1 Operationalization

Code	Survey questions under operationalized concepts	Hypotheses
	Demography	
Age	What year were you born?	
Gender	What gender do you identify as?	
Educational level	What is the highest level of education you have completed?	
Occupation	What is your main occupation?	
Resident country	Do you live in Sweden?	
	Perceived benefits	
PB1	I appreciate when streaming services know my personal preferences and adapt the service to me.	H1
PB2	I am okay with sharing my personal data with a streaming service if it helps me find what I am searching for faster.	H1
PB3	I am okay with sharing my personal data with a streaming service if it gives me a better user experience.	H1
PB4	I am okay with sharing my personal data with a streaming service to receive recommendations on content relevant to me.	H1
	Perceived risks	
PR1	I see risks in sharing my personal data online.	H2
PR2	I refrain from sharing my personal data if I feel that the risks outweigh the benefits of doing so.	H2
PR3	I am worried about the wrong person getting hold of my personal data when I share it with companies online.	H2
PR4	I can feel monitored when I receive personalized recommendations, as if the company knows too much about me.	H2
PR5	Personalized recommendations complicate my ability to discover new content that I do not usually consume.	H2
PR6	I find suggestions for content that is not relevant to me annoying.	H2
	Trust	
TU1	I am confident in sharing my personal data with companies that I trust.	H3
TU2	Overall, I have more trust in Swedish companies than in foreign companies.	H3
	Transparency	
TY1	It is important for me that companies clearly communicate how they collect my personal data.	H4
TY2	I usually read the privacy policy before sharing my personal data with companies.	H4
TY3	I am okay with sharing my personal data if companies are clear about how they use it.	H4

	Context	
	Have you used Netflix/Spotify any time in the last year?	
CO1	I am aware of how Netflix/Spotify use my personal data.	H5
CO2	I am okay with sharing my personal data with Netflix/Spotify.	H5
CO3	I see risks in sharing my personal data with Netflix/Spotify.	H5
CO4	I trust that Netflix/Spotify process my personal data correctly.	H5
CO5	I am in favor of receiving personalized recommendations from Netflix/Spotify based on my personal data.	H5
	Willingness to receive personalized communication based on personal information	
WR1	If I experience benefits from personalized recommendations, I want to continue receiving them.	H1
WR2	If I experience risks from personalized recommendations, I do not want to continue receiving them.	H2
WR3	If I have a high level of trust in a company, I am positive to receive personalized recommendations based on my personal data.	H3
WR4	If a company is transparent about how they use my personal data, I am positive to receive personalized recommendations based on my personal data.	H4
WR5	The nature of the company affects my willingness to receive personalized recommendations based on my personal data.	H5
WR6	I prefer personalized recommendations over searching for what I like myself.	