Big Brother is Watching You –

The Impact of Consumer Understanding and Awareness of Data-Based Personalization on Consumer Attitude

Student Name:	Pascale Bastian
Date:	24th June 2017
University:	University of Twente
Faculty	Faculty of Behavioral Management and Social Sciences
Master	Communication Studies, Marketing Communications
1st Supervisor:	Dr. Thomas van Rompay
2nd Supervisor:	Dr. Suzanne Janssen

Abstract

Smart technologies spread into more aspects of consumers' daily lives, leading to the creation of massive data sets. These give real-time insights into actual consumer behavior, which results in an increasing personalization of the consumer web experience, thereby creating a privacy paradox: Consumers are highly concerned about online privacy, but also do not invest the time to carefully read and understand companies' privacy statements. To reduce this privacy paradox consumer understanding and awareness of data-driven personalization needs to be facilitated. The facilitation of consumer reading through visual eye cues could increase consumer awareness, whereas clear and comprehensible privacy statements could enhance consumer understanding.

The aim of this study was to test whether the facilitation of consumer understanding and awareness of data collection and sharing practices affects consumer attitude towards data-driven personalization and companies engaging in it.

A 3 (personalization: no, medium, high) x 2 (watching eyes: eyes present, no eyes) factorial design was used. The dependent variables were app data safety, data sharing attitude, company trustworthiness, company likeability and intention to download the app. An online survey was conducted with 151 participants. The stimulus was a fitness app, with the app permissions functioning as experimental manipulation.

The results show that increasing levels of personalization have a significant negative effect on consumer attitude towards app data safety, data sharing attitude, company trustworthiness and likeability. There was no facilitation effect of the watching eyes.

This study shows that consumer understanding of data-driven personalization can be increased through the clear and comprehensible declarations of data usage. This enables consumers to recognize privacy threats more easily and protect their data. Hence, for a company to be successful in the long run more consumer friendly data collection models are necessary, to enable consumers to actively decide which data they want to share and with whom.

Keywords: personalization, big data, online privacy, health apps, watching eyes

1 Introduction

George Orwell's (1949) quote "Big Brother is Watching You" to describe an authoritarian regime that is spying on and manipulating its citizens, may as well be used to describe today's digitalized society (p.3). The proliferation of smart technologies and increasing interconnectedness between people through these smart devices has revolutionized the ability to communicate, share and access information (Al-Khouri, 2013; Tene & Polonetsky, 2012). This leads to the creation of massive amounts of data, as every web interaction whether it is a search query on the internet or using an app, leaves behind a trace of data. Moreover, if analyzed these massive sets of data can give companies real-time insights on consumer preferences, desires and actual behaviors. For example, these insights may be gained from consumer online transactions, social networking activities, browsing behavior or app usage (Erevelles, Fukawa, & Swayne, 2016). Thereby, this so called big data, allows for a much deeper understanding of consumer behavior, enabling real-time personalization of the consumer web experience (Stoicescu, 2015). Real-time personalization includes personalized content according to consumer interests or political affiliation, targeted advertisements or even individualized website designs based on consumer preferences, age or gender (Pariser, 2011).

The personalization of consumer web experience definitely has its benefits in terms of convenience and relevance, as these processes filter the vast amounts of information available online to present consumers with more relevant content (O'Connor, 2007; Teltzrow & Kobsa, 2004). However, there are also massive implications in terms of consumer privacy as most consumers are unaware of companies' data collection and data sharing practices (Labrecque, vor dem Esche, Mathwick, Novak, & Hofacker, 2013). The problem with data-driven personalization is that, on the one hand, companies lack in providing comprehensive information about their data collection practices to consumers, whereas consumers, on the other hand, do not devote the time to read and understand privacy statements provided by companies. According to a study by Quint and Rogers (2015), the majority of consumers want more information about what data companies collect about them as well as have greater control over the amount of data being collected. However, at the same time they do not invest time to read and understand the provided privacy statements or app permissions (Meinert, Peterson, Criswell, & Crossland, 2006). This creates a privacy paradox as consumers are concerned about privacy, but on the other hand do not act to protect their privacy by investing time and energy in understanding companies' data collection and sharing practices. Therefore, to eliminate this privacy paradox consumer understanding and awareness of data-driven personalization needs to be increased, so that consumers can efficiently protect their privacy and exert more control over their personal data in an online environment. For example, consumer understanding can be increased by presenting the information in a more simplified language, whereas consumer awareness may be increased by facilitating consumer reading through the implementation of visual cues.

Several studies that were conducted on online privacy and consumer awareness of privacy issues found that consumers are concerned about how companies collect and use personal data. Moreover, these studies showed that there is a lack in consumer awareness and understanding of these data collection and sharing practices that happen beyond consumer control. Thus, to allow consumers to actively protect their privacy and control companies' usage of personal data, there is a need to increase consumer awareness and understanding of data collection practices that drive personalization. Therefore, the aim of this study is to test whether the facilitation of consumer understanding and awareness of a company`s exact data collection and data sharing practices affects consumer attitude towards the company and data-driven personalization. An experiment was conducted with different levels of personalization and a visual eye cue to facilitate consumer reading and understanding of data-driven personalization. The central question guiding this research was as follows: *"How does the facilitation of consumer understanding and awareness of data-driven personalization affect consumer attitude towards data-driven personalization and companies engaging in it?*

2 Theoretical Framework

2.1 What is Big Data?

The term Big Data refers to enormous and highly complex data sets, structured and unstructured, that if analyzed can reveal important underlying and highly valuable relationships for the global economy. It provides detailed information and rich insights that can drive product innovation, productivity, efficiency and growth of the economy (Tene & Polonetsky, 2012).

There are three key dimensions of Big Data: volume, velocity and variety. Volume refers to the vast amounts of data which are currently measured in petabytes, where 1 petabyte is equivalent to roughly 20 million traditional filing cabinets of text. The amount of data collected is expected to double in size every two years and estimated to reach 44 zettabytes by 2020 (Erevelles et al., 2016). The second dimension velocity, refers to the relentless speed of data creation that allows real-time insights on all consumer activities (Stoicescu, 2015). And finally, the third dimension is variety which denotes the diversity and richness of the multiple types of data including structured data such as scanner or sensor data as well as unstructured data such as videos, blogs, audio-recordings and text messages (Erevelles et al., 2016; Stoicescu, 2015). Furthermore, there are two additional dimensions that play an important role in structuring the data. Veracity highlights the need to be aware of data quality since not all big data about consumers is accurate and value refers to the importance of eliminating irrelevant and unimportant data so that the remaining information is useful (Erevelles et al., 2016).

2.2 Big Data and Consumer Analytics

Traditional consumer research methods were only capable of capturing the behavioral intentions of consumers but not their actual behavior. With vast amounts of information now being available on actual consumer behavior and experiences this offers the opportunity of revolutionizing our understanding of consumer behavior (Stoicescu, 2015). Companies can gain real-time insights on consumer preferences, transactions, social network activities, search histories and any complaints or inquiries as well as the geographical mobility of their customers. For example, apps like

Foursquare or Loopt analyze a consumer's geographical location as well as personal interests to target consumers with personalized ads while they are shopping in-store. Moreover, when combining all this information with physiological data from wearables or health apps and historical purchase data, this provides companies with previously unknown insights that reveal an enhanced level of understanding of consumers (Erevelles et al., 2016). For example, if information about physical activity or eating habits falls into the hands of an insurance company, then this company can base their monthly rate for a customer on his or her overall health. Thereby, big data allows for a much deeper understanding of consumers enabling real-time individualization of marketing activities, see Figure 1 (Stoicescu, 2015)**.** According to a study by eMarketer.com (2013), 65% of US marketing agencies say that Big Data helps them gain greater insight into customer experiences across all media types and craft effective marketing strategies.

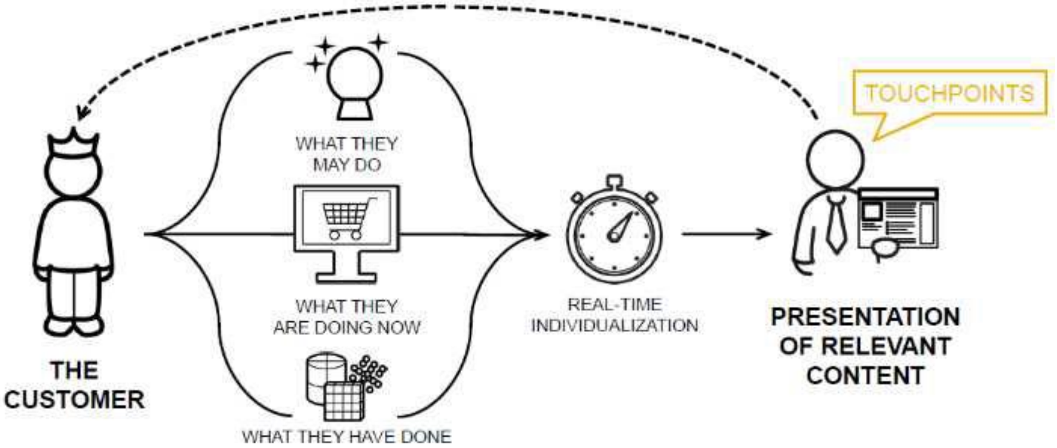


Figure 1 The process of real-time individualization (Stoicescu 2015).

The extraction of hidden patterns, correlations and other insights from large amounts of data is called big data analytics (SAS, 2016a). Big data analytics includes several types of technologies that help to manage, extract and analyze the most valuable information for a set purpose. According to the data scientist Bertolucci (2013), there are three stages of data analytics. The first step is descriptive analytics which is used to summarize large amounts of historical data into smaller more useful sets of information (Bertolucci, 2013). The second step is predictive analytics where several statistical, modeling, data mining and machine learning techniques are used to analyze historical data from descriptive analytics and recent data to make predictions about the future (Predictive Analytics Today, 2014). For example, the music streaming service Spotify uses predictive analytics to send their users personalized recommendations of other artists or bands based on their current playlist. The next step is prescriptive analytics which essentially uses the information from descriptive and predictive analytics to make recommendations about future actions based on desired outcomes (Bertolucci, 2013; Oracle, 2012). One example of prescriptive analytics is Google's new self-driving car as it makes multiple decisions about what to do next when standing at an intersection based on calculated future outcomes (van Rijmenam, 2014). These algorithms can act as the basis for developing strategies that drive product innovation, productivity, efficiency and economic growth.

The increasing personalization of the consumer web experience is mostly based on the analysis of historical and recent consumer activity, thereby using descriptive and prescriptive analytics. Companies move away from broad customer segments with messages appealing to the masses to a more granular level with highly personalized messages customers can identify themselves with. This is done with the help of individual consumer profiles which include information about online purchases, online behavior, social media, direct communications, email interactions and interests. Thereby, allowing for the creation of a detailed profile which may even include the type of persuasion a consumer is most susceptible to (Pariser, 2011). Moreover, big data allows marketers to track their customer segments as they evolve over time and always keep up to date with changing customer interests (Reachforce, 2015). This improves the effectiveness of marketing activities as it enables marketers to target individual customers with products or services that are tailored to their specific needs at the time the need arises. However, what are the implications of this personalization for consumer privacy?

2.3 Implications of Personalization for Consumers & Online Privacy

Personalization is a customer-oriented marketing strategy that is based on data-driven insights about individual consumers with the aim of targeting them with personalized content in real-time (Aguirre, Mahr, Grewal, Ruyter, & Wetzels, 2015). This brings convenience and more value for consumers, as they are presented with personalized content or advertisements that meet their needs instead of having to waste endless time and efforts searching the web for relevant content (O'Connor, 2007; Teltzrow & Kobsa, 2004).

When looking at the vast amounts of information that is available online, the process of filtering and personalizing information of course has its benefits for the consumer in terms of relevance and convenience (Pariser, 2011). However, there are also negative implications, especially concerning consumer privacy. Today’s social web empowers people to gain knowledge, follow their interests and ideals, interact socially, share their opinions and discover new things. But this empowerment in the age of big data often comes at the price of privacy (Labrecque et al., 2013). Thereby, creating a technological paradox that is empowerment of consumers vs. enslavement of consumers to marketers, because individuals share more personal information online, often without being aware of how that data is used by third parties (Mick & Fournier, 1998).

The problem with data-driven personalization on the one hand is that much of the data collection and data sharing is happening without consumer awareness and control. A study by Malhotra, Kim, and Agarwal (2004) found that consumers were less concerned about data privacy when they were made aware of the data collection and had direct control over their personal information by giving their informed consent. However, much of the personal information is collected using covert information collection strategies that unobtrusively gather consumer data by tracking their browsing behavior (Aguirre et al., 2015; O'Connor, 2007). Mostly this is done via the placement of cookies, which are text files that are stored on a consumer’s computer and offered back to the website once the consumer revisits. Some types of cookies are required for websites to work effectively, for example, to store items in a virtual shopping cart. However, alongside these required cookies, cookies from third-party organizations are often also placed on the user’s hard drive and used to track consumer behavior across websites for marketing purposes (Miyazaki, 2008). Hence, the information gathered through cookies can be used for real-time personalization of the consumer web experience. Most companies declare their collection of personal information in cookie disclaimers or more extensively in privacy statements (Morey, Forbath, & Schoop, 2015). However, these privacy statements are quite long most of the time and the use of incomprehensible or vague language, makes it impossible for consumer to understand (Marotta-Wurgler, 2005; McDonald & Cranor, 2008).

So, even if companies declare their collection of personal information in cookie disclaimers, privacy statements or app permissions to give the consumer the feeling of control, the consumer is in fact still left in the dark about the company's true data practices. To facilitate consumer understanding and awareness of data-driven personalization and the resulting privacy issues, privacy statements need to be more clear and comprehensive as well as stand out to attract consumers awareness (Luzak, 2014).

2.4 Consumer Attitude about Personalization

Consumers have a poor understanding of the complex processes involved in the collection and aggregation of personal data (Quint & Rogers, 2015). Furthermore, consumers are becoming increasingly concerned about companies sharing personal information with third parties (Strong, 2013). According to Quint and Rogers (2015), about 90% of consumers are at least somewhat concerned about data privacy and how companies are using personal data. Furthermore, the majority of consumers, 85%, want to be more informed about the type of personal information companies collect and 86% want to have more control over the information stored by companies. One reason for that is that many consumers, 75%, are concerned that sharing their personal data makes them targets for marketing campaigns (Quint & Rogers, 2015). In another study, it was found that only about one third of consumers have a positive attitude towards companies using consumer data for personalization. Whereas, the majority of consumers, 69%, have are more negative attitude towards companies using personal data for personalization and even experience a feeling of creepiness when confronted with personalized advertisements (Strong, 2013). This creepiness factor occurs when consumers get a sense of violation, because an advertisements is too personal and/or based on data consumers did not agree to give (Barnard, 2014). To test these assumptions the following hypotheses were derived:

H1. Consumers experiencing a high level of personalization will have a more negative attitude towards app data safety and data sharing than consumers experiencing a low level of personalization.

However, consumer concerns about data-driven personalization and data sharing, as well as their desire to have more control over personal data do not prohibit them from wanting to share. Quite the contrary, consumers are willing to share personal information when brands use this information to create value for them. Hereby, brand trust plays an important role in influencing consumers' willingness to share personal information (Gigya, 2015). According to Quint and Rogers (2015) roughly 75% of consumers are willing to share data with brands they trust and about 80% of consumers would even share a non-required piece of data in exchange for reward points or other incentives. Furthermore, women are 10% more likely to share their email address for offers (Gigya, 2015). However, for consumers to be willing to share their data with a company they need reassurance that the information will only be used by that company (EY, 2013; Quint & Rogers, 2015).

H2. Consumers experiencing a high level of personalization will have a more negative attitude towards company likeability and company trustworthiness than consumers experiencing a low level of personalization.

2.5 Consumer Awareness of Personalization

Most consumers are aware that companies are collecting personal information about them, but are poorly informed about the type of data collected and how it is used (Morey et al., 2015). For example, most consumers know that companies such as Google, Facebook and Amazon collect personal data to target them with personalized advertisements, yet they are not aware of the extent of data collection and have no control over the process. As mentioned in the previous section, most of this web tracking is done by the placement of cookies. A study conducted by Jensen, Potts and Jensen (2005) revealed that even though 90.3% of their participants claimed to have knowledge of cookies, only about 15.5% were able to demonstrate simple cookie knowledge. Thus, most consumers are probably unaware of the extent these cookies are used to personalize their web experience and their personal information often being sold off to unknown third parties, despite visible cookie disclaimers on most websites (Hill, 2015).

The mobile counterpart to desktop cookie disclaimers are app permissions. These app permissions have to be accepted by the user directly after downloading and immediately before being able to use the app. Similar to cookies, most users do not fully understand the content of app permissions and by now they have become used to them, thus often accept them mindlessly (Chia, Yamamoto, & Asokan, 2012). This mindless behavior is more concerning on mobile devices than on desktop, because mobile devices are much more intimate and personal to the consumer. Most people carry their mobile device around wherever they go and use it for more sensitive activities such as personal communication or banking (Meng, Ding, Chung, Han, & Lee, 2016). A study on mobile apps found that 73% of Android Apps share highly sensitive information such as personal email address with third parties. Furthermore, about 47% of iOS apps share location data with third party organizations (Zang, Dummit, Graves, Lisker, & Sweeney, 2015). Thus, by simply accepting app permissions regardless of their content consumers put their privacy at risk and willingly give their personal data to unknown third parties. These third parties can then use the information to target consumers with personalized advertisements or individual content. Therefore, it will be interesting to see whether the facilitation of consumer understanding and awareness of a company's data collection practices affects consumer willingness to download an app. The following hypothesis was derived:

H3. Consumers experiencing a low level of personalization will have a higher intention to download the app than consumers experiencing a high level of personalization.

From the consumer perspective, the problem with data-driven personalization is that consumers do not invest the time to read and understand privacy statements to be able to protect their privacy. Therefore, they are poorly informed about the type of personal information companies collect and how this data is used beyond the originally intended purpose (Morey et al., 2015). So, if people are so concerned about their privacy online why do they not read privacy statements then? Hereby, it is important explore the possible benefits and costs consumers get out of reading privacy statements. One major costs is definitely time, privacy statements are very long and written in a more complicated language, thus it takes a long time to read and understand the information presented (McDonald & Cranor, 2008). In their study Mc Donald and Cranor (2008) estimated that it would take each consumers about 244 hours on average to read the privacy statements of each website visited in one year. Another reason, might be that the information may not be relevant to the consumer in the moment of presentation. For example, when a consumer is searching for a specific product online, he or she does not want to spend an excessive amount of time reading a privacy statement before purchasing that product. Furthermore, consumer's cost-benefit tradeoffs regarding the sharing of their personal information are often context-specific, meaning that if they like a specific product or service, information about possible privacy threats is often ignored (Marotta-Wurgler, 2005). This can be referred to as personal interest in a certain product or service, as this variable highly depends on the person it cannot be manipulated in an experiment. Therefore, personal interest will be defined as a moderator.

If it is true that consumer's make context-specific decisions when it comes to sharing of personal information, then it is likely that there will be an interaction between the level of personalization and personal health interest of the consumer. Therefore, the following hypotheses were derived:

H4. The effects of level of personalization on consumer attitude towards app data safety and data sharing will be stronger for consumers with a low health interest than for consumers with a high health interest.

H5. The effects of level of personalization on consumer attitude towards company likeability and company trustworthiness will be stronger for consumers with a low health interest than for consumers with a high health interest.

2.6 Level of Awareness – The Feeling of Being Watched?

When it comes to consumer attitude about data-driven personalization, the level of awareness consumers have about these processes plays an important role. Essentially, attitudes are evaluative judgments of how good or bad a consumer finds a particular issue, place or product (Kardes, Cline, & Cronley, 2011). So, if one is not aware of something, then how could one judge it? Therefore, it would be interesting to see whether the level awareness of consumers and the resulting attitude could be enhanced by subtle visual cues in the environment. There is increasing agreement among behavioral scientists that human behavior can indeed be significantly influenced by factors relating to the situation or context in which the information is presented (Dolan, Hallsworth, Halpern, King, Metcalfe & Vlaev, 2012). This is due to the fact that the human brain relies heavily on fast, simple and non-conscious "rules-of-thumb" while scanning the surroundings to evaluate the situation and form a subsequent decision on how to behave. Thus, any change to the environment no matter how insignificant it may seem can lead to a large change in behavior (Nettle, Nott, & Bateson, 2012).

An example of how subtle cues in the environment can affect human behavior is the 'watching eyes' effect. A series of studies including controlled laboratory experiments and real-life settings have shown that portraying images of stylized-eyes induced people to behave more pro-socially, even in completely anonymous situations. Some examples of real-life settings include donations to charity, following of recycling rules, reduced littering as well as bicycle thefts (Bateson, Callow, Holmes, Redmond Roche, & Nettle, 2013; Haley & Fessler, 2005; Nettle et al., 2012; Pfattheicher & Keller, 2015). According to Bateson et al. (2013), this pro-social behavior can be linked to reputation-based partner choice models which suggest that people behave pro-socially, even when no immediate returns can be expected, to invest in their social reputation. This is especially true for individuals with a strong public self-awareness as they are particularly concerned about how others see them and their social reputation (Pfattheicher & Keller, 2015). Pfattheicher and Keller (2015) suggest in their study that the degree of public self-awareness in individuals might act a as a moderator of the watching eyes phenomenon.

In light of these findings, it might be interesting to see whether the watching eyes effect also has an impact on human behavior in other scenarios that are not directly linked to the domain of antisocial or prosocial tendencies. For example, whether the watching eyes phenomenon could be used to raise consumer awareness of data-driven personalization such as encouraging people to actively read app permissions or cookie disclaimers. According to Haley and Fessler (2005), humans have an innate reaction to faces, hereby the eyes as stimulus play an essential role. Over the course of the evolution eyes facing in one's direction have been a clear indicator that attention is being direct towards oneself. Thus, humans have evolved to respond to all eye-like stimuli, even if they are just representations and not actual eyes (Haley & Fessler, 2005; Nettle et al., 2012).

Several studies have demonstrated, that the presence of watching eyes increases consumer compliance with a request. Therefore, this study will explore whether app permissions accompanied by watching eyes facilitate consumer awareness when it comes to reading and processing the information provided in the app permissions. To test these assumptions the following hypotheses were derived:

H6. If it is true that watching eyes increase consumer awareness and facilitate consumer reading of app permissions, then it is expected that the above-mentioned effects of level of personalization in H1 and H2 will be stronger in the presence of a visual eye cue.

3 Methodology

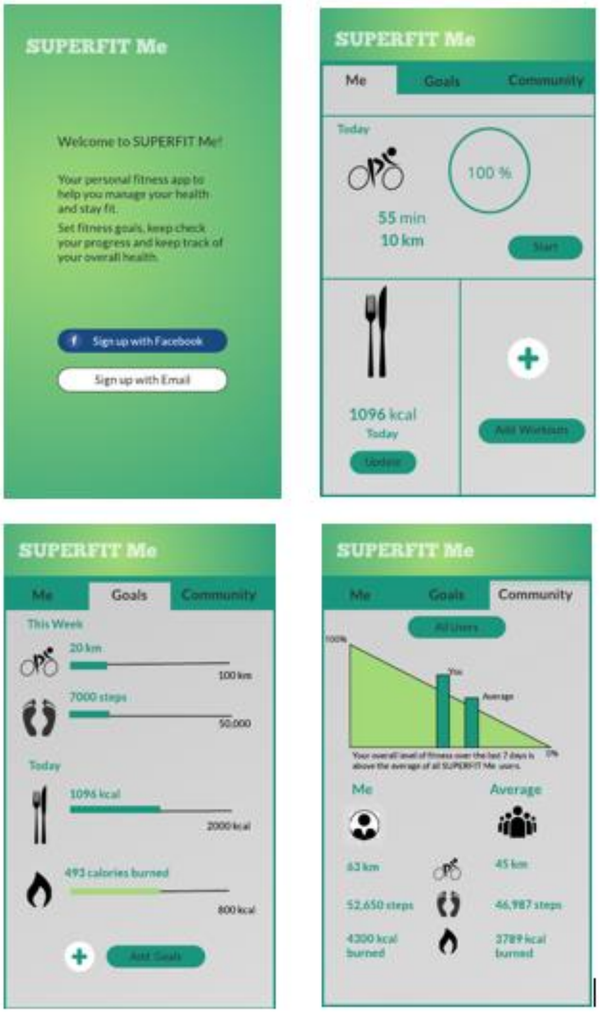
3.1 Design

To empirically test the hypotheses stated in the previous section a 2 (watching eyes: eyes present vs. no eyes present) x 3 (personalization: no, medium, high) factorial design was implemented. The design consists of two independent variables, level of personalization and watching eyes as well as the moderator personal health interest. The first independent variable, level of personalization, consists of three experimental conditions: 1) no personalization 2) medium personalization and 3) high personalization. The second independent variable, watching eyes, consists of two experimental groups: 1) watching eyes present and 2) no eyes present. The dependent variables are consumer attitude towards app data safety, data sharing attitude, company trustworthiness, company likeability and intention to download the app. The experimental manipulation took place in the form of app permissions being presented to the participants in an app testing situation.

The 2x3 factorial design was chosen because it shows whether consumer awareness and understanding of a company's data practices affects consumer attitude towards data-driven personalization as well as towards companies that engage in it. Furthermore, it will show whether consumer reading of app permissions can be facilitated by the presence of watching eyes.

3.2 Stimulus Material and Manipulations

In order to study how different levels of personalization affect consumer attitude about big data practices, a relevant and realistic setting had to be chosen for the experiment. Based on the theoretical framework, consumers often behave mindlessly when accepting app permissions on their smartphone, this provides an ideal precondition for testing how consumer awareness can be influenced in these situations. Another important element of the context is the type of information required from consumers. According to Malhotra et al. (2004), financial and health data are considered as more sensitive information by consumers, whereas lifestyle or shopping characteristics are viewed as less sensitive. For this study, a fitness app context was chosen, because fitness apps are becoming more popular among consumers due to the increasing availability of fitness trackers and wearables (Statista, 2015). Moreover, fitness apps mostly collect highly sensitive consumer data and are considered to be unsafe in terms of consumer privacy (Peterson, 2014).



Welcome to SUPERFIT Me!

Your personal fitness app to help you manage your health and stay fit.

Set fitness goals, keep check your progress and keep track of your overall health.

Your overall level of fitness over the last 7 days is above the average of all SUPERFIT Me users.

Figure 2 The prototype of the fitness app 'SuperFit Me'

To create a realistic experimental setting a prototype for the app was designed. In line with the fitness context, the app was called ‘SuperFit Me’ to give participants an immediate idea what the app is about. Also, at the time of the study there was currently no fitness app on the market with a similar name. The design of the app was based on popular fitness apps such as S Health and Runkeeper to make the app more realistic and give participants a feeling of familiarity. Moreover, to give participants a better idea of the app functionalities they were presented with 4 screens, similar to the app preview in an app store. Thus, the app design entailed a Welcome Screen, a Me Screen, a Goal Screen and a Community Screen, as shown in Figure 2. Additionally, one more screen was created for the experimental manipulation, displaying the app permissions. This screen will be discussed in the next section.

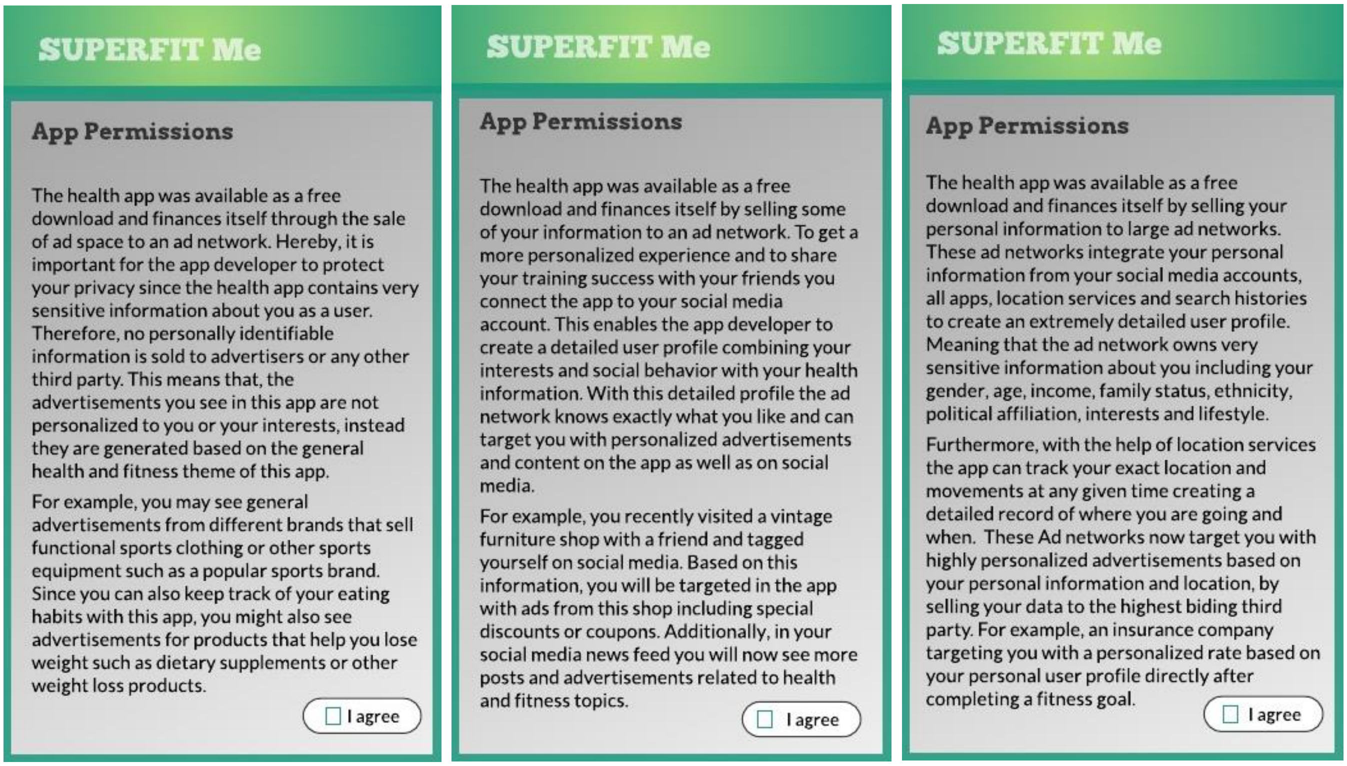
On the ‘Welcome Screen’ the participants were presented with a welcome note and the option to login either via Facebook or email address. These login options were chosen, because they are provided by most apps on the market, thus consumers are used to them. The second screen, ‘Me’, gives participants an overview of their recent training sessions as well as eating habits, with the option of adding additional workouts or other health measures. On the third screen, ‘Goals’, participants can view their training or dietary goals and the timeframe for achieving these goals, enabling them to keep a close eye on their level of fitness and personal development. The fourth screen, ‘Community’, gives participants the opportunity to compare their performance to friends or even the entire ‘SuperFit Me’ community. These four screens give participants a feeling for the app and its functions, thus making the experimental setting more realistic and adding to the reliability of the study.

Manipulation Check. As mentioned in the previous section, the app permissions functioned as the experimental manipulation, and differed for each condition in their level of personalization as well as whether watching eyes were present or not. Therefore, for the first independent variable three app permissions scenarios had to be written ranging from no personalization to extreme personalization. To ensure the objectivity of the scenarios and relevance of the chosen context a pretest was designed (Appendix C). Based on the fitness app context mentioned in the previous section 5 app permission scenarios were written ranging from no personalization to extreme personalization. In the pretest, the respondents were first presented with a short description of the app functions and then were asked to rate each scenario on how realistic, invasive and disturbing they felt it was. The five scenarios were presented randomly to the respondents. For this measurement, a 7-point semantic differential scale was used with the following items: Realistic – Unrealistic, Invasive – Protective, and Disturbing – Acceptable.

The online questionnaire was administered to the respondents via Qualtrics (www.qualtrics.com) and was filled in by a total of 15 respondents. The results show that all 5 scenarios were rated as realistic with the third scenario (M=2,13) being the most realistic and the fifth scenario (M=3,00) being the least realistic. However, the scenarios clearly differ in the level of invasiveness and disturbingness. Based on the results scenario 1, 3 and 5 were chosen for the experiment, because they provide a good range in terms of realism, invasiveness and disturbingness. Scenario 1 describes the lower end of the spectrum as it is rated as realistic (M= 2,73), very protective (M= 5,27) and acceptable (M= 5,67) by the respondents, thus it provides a good neutral base for comparison. The medium level of personalization scenario is scenario 3, as it is rated as somehow invasive (M=2,60) and disturbing (M=3,33), making it a good contrast to scenario 1, but not as shocking as scenario 5. Scenario 5 was considered as very invasive (M=1,60) and disturbing (2,07) by respondents, thus it constitutes a very extreme situation that may evoke a more protective response from participants in terms of privacy violation. Thus, scenario 1 is the no personalization condition, scenario 3 is the medium personalization condition and scenario 5 is the extreme personalization condition.

Level of Personalization. As mentioned above the first independent variable, level of personalization, entails three experimental conditions: 1) no personalization 2) medium personalization and 3) high personalization. Each of these conditions consists of a scenario describing app permissions which differ in the level of data-based personalization. The scenarios were presented in the framework of the fitness app ‘SuperFit Me’. Based on the pretest, the following scenarios were chosen: 1) Scenario 1 = no personalization 2) Scenario 3 = medium personalization and 3) Scenario 5 = extreme personalization. In the following sections for clarification these scenarios were referred to as no personalization, medium personalization and high personalization. All three written scenarios are portrayed below in Figure 3. To control for any bias, the scenarios were written in an objective tonality and with the intention of presenting the information without any judgment. Furthermore, all three scenarios have the same length and describe the same context. The only variation in the scenarios is the level of personalization to control for contextual bias and increase the reliability of the study.

Figure 3. App Permissions with the experimental conditions: no Personalization (left), medium personalization (middle) and high personalization (right).



SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself through the sale of ad space to an ad network. Hereby, it is important for the app developer to protect your privacy since the health app contains very sensitive information about you as a user. Therefore, no personally identifiable information is sold to advertisers or any other third party. This means that, the advertisements you see in this app are not personalized to you or your interests, instead they are generated based on the general health and fitness theme of this app.

For example, you may see general advertisements from different brands that sell functional sports clothing or other sports equipment such as a popular sports brand. Since you can also keep track of your eating habits with this app, you might also see advertisements for products that help you lose weight such as dietary supplements or other weight loss products.

□ I agree

SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself by selling some of your information to an ad network. To get a more personalized experience and to share your training success with your friends you connect the app to your social media account. This enables the app developer to create a detailed user profile combining your interests and social behavior with your health information. With this detailed profile the ad network knows exactly what you like and can target you with personalized advertisements and content on the app as well as on social media.

For example, you recently visited a vintage furniture shop with a friend and tagged yourself on social media. Based on this information, you will be targeted in the app with ads from this shop including special discounts or coupons. Additionally, in your social media news feed you will now see more posts and advertisements related to health and fitness topics.

□ I agree

SUPERFIT Me

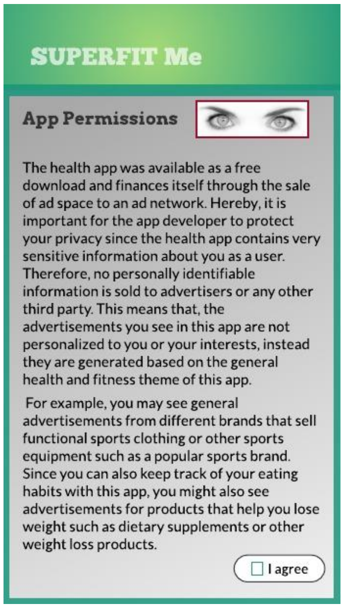
App Permissions

The health app was available as a free download and finances itself by selling your personal information to large ad networks. These ad networks integrate your personal information from your social media accounts, all apps, location services and search histories to create an extremely detailed user profile. Meaning that the ad network owns very sensitive information about you including your gender, age, income, family status, ethnicity, political affiliation, interests and lifestyle.

Furthermore, with the help of location services the app can track your exact location and movements at any given time creating a detailed record of where you are going and when. These Ad networks now target you with highly personalized advertisements based on your personal information and location, by selling your data to the highest biding third party. For example, an insurance company targeting you with a personalized rate based on your personal user profile directly after completing a fitness goal.

□ I agree

Watching Eyes. The second independent variable watching eyes consists of two experimental conditions: 1) stylized eyes present 2) no eyes present. Each participant was randomly assigned to one of the two conditions. As shown in Figure 4, for participants exposed to the visual eye cue, the app permissions screen included a pair of stylized eyes in the top right corner next to the heading ‘App Permissions’. For the subtle eye cue a realistic drawing of eyes was chosen, as it creates a more effective feeling of being watched than an abstract eye cue would. Furthermore, the eye image was surrounded by a red frame, to further intensify the feeling of being watched and to subconsciously raise consumer awareness. In all three personalization vs. eyes present conditions, the size of the watching eyes was 2.4 x 0.8 cm and they were positioned in the same place next to the heading ‘App Permissions’. In the no eyes condition, there was no eye cue present and participants were exposed to the app permissions, as shown in Figure 3 in the previous section.



SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself through the sale of ad space to an ad network. Hereby, it is important for the app developer to protect your privacy since the health app contains very sensitive information about you as a user. Therefore, no personally identifiable information is sold to advertisers or any other third party. This means that, the advertisements you see in this app are not personalized to you or your interests, instead they are generated based on the general health and fitness theme of this app.

For example, you may see general advertisements from different brands that sell functional sports clothing or other sports equipment such as a popular sports brand. Since you can also keep track of your eating habits with this app, you might also see advertisements for products that help you lose weight such as dietary supplements or other weight loss products.

Figure 4 App Permissions screen with Watching Eyes.

3.3 Procedure

The experiment was administered in the form of a quantitative online survey, as this provides a time and cost-effective data collection method as well as adding to the anonymous online context. Due to is anonymity, this data collection method constitutes a normal online usage experience for the participants. The questionnaire was provided as desktop and optimized mobile version, as mobile devices such as smartphones or tablets are the primary source of app downloads and accounted for 51.3% of internet usage worldwide in 2016 (StatCounter, 2016). The questionnaire was designed with the Qualtrics Survey Software tool (www.qualtrics.com). The data collection took place from 25.01.2017 till 20.02.2017.

In the introduction of the questionnaire participants were informed that the purpose of this survey was to gather their feedback on the usability and design of a new fitness app (see Appendix A). This would be done by showing them screenshots of the app functions and then asking them to download an app prototype for testing. The participants were led to believe that the study was about testing a new fitness app, to make the setting as neutral and realistic as possible, thereby reducing any type of priming or bias that could falsify the results. Moreover, participants were not provided with an informed consent form as this could have primed them with the concept of giving consent, which could falsify the results. However, it was communicated in the introduction that participants will have the chance to win an Amazon voucher worth 20€. This was not only implemented as an incentive, but also to test whether participant's willingness to share personal data differed between experimental groups.

After the introduction, participants were presented with a normal usage situation for the app, describing the app functionalities such as the use location services to track speed and distance of bike rides. This was done to give participants a better feel for the app and the testing situation. Since this survey was conducted online, it was not possible to control for the surroundings or cognitive state of participants, thus it was important to induce the correct mindset by providing additional information that sets the scene.

Then, participants were exposed to the app screens described in the previous section. The first screen was the Welcome Screen with the login options, here participants were asked to specify which login option they prefer. This question was used as a filler question, and to get participants attention after having to read so much information by requiring an action. Furthermore, as participants were not able to interact with the app, each app screen was presented with a short description of the functionalities. The other screens as described in the stimulus section did not include any questions and were presented in following order: Me, Goals and Community. After giving participants a feel for the app and its functionalities, they were presented randomly with 1 of 6 manipulated app permission screens, the experimental condition. Then, participants had to fill out a questionnaire about their attitude towards the app, the company behind the app and about big data, personalization and data sharing. After the questions about the app and the company behind the app, participants were asked whether they wanted to download the app now. This question was added to see their general attitude towards the app and give another measure of comparison between the experimental groups.

However, as this study was not actually about testing and downloading the app, participants were led to a screen informing them about the actual background of the study, namely assessing consumer attitude towards data-driven personalization and companies engaging in it. This was followed up by another section of questions to test participant's knowledge and attitude towards data-driven personalization and data sharing. Furthermore, to gather some qualitative data participants were asked to state their thoughts and main concerns about the increasing personalization and data sharing. Finally, the questionnaire was rounded up by two demographics questions about age and gender, to see if there are any significant differences relevant to this study. One final test was added, if participants wanted to win the Amazon voucher, they should provide their personal email. It would be interesting to see if participants differ in their willingness to give their personal email address depending on the assigned experimental condition.

3.4 Questionnaire

The survey consists of eight constructs: general app evaluation, app data safety, personal health interest, company trustworthiness, company likeability, data sharing attitude, privacy concern and general big data knowledge. The dependent variable intention to download the app was measured by one closed question. Apart from these constructs, two demographics questions about participants age and gender were added as well as two open questions about participant's general thoughts and main concerns about data-driven personalization. To measure participant's attitude towards the different constructs a 5-point Likert-scale was used. The scale ranged from strongly agree (1) to strongly disagree (5). A scale developed by Malhotra et al. (2004) to study user's information privacy concerns was used to guide the creation of statements and constructs for this study. To measure the internal consistency between the items of each construct Cronbach's alpha was conducted.

General App Evaluation. The construct general app evaluation consisted of 3 statements measuring the usability and the design of the app. This construct was added because the design and usability are important factors in determining consumer attitude about products (Kardes et al., 2011). Therefore, general app evaluation will function as a control variable. The Cronbach's alpha was .62, which means an acceptable reliability. An example item of this construct was 'I like the design of the app'.

Personal Health Interest. The construct personal health interest consisted of 4 items measuring participant's interest in health apps and leading a healthy lifestyle. According to a study by Marotta-Wurgler (2005), consumer's privacy concerns are often moderated by their interest in a specific product or service. Thus, in this study personal health interest will act as a moderator variable. The Cronbach's alpha measured was 0.72, represent a good reliability for this construct. An example item was 'This app can help me lead a healthy lifestyle'.

App Data Safety. The construct app data safety consisted of 4 statements measuring participant's feeling of data safety that the app provides. The Cronbach's alpha for this construct was originally .84, which already indicates a very good reliability. However, by removing one item the Cronbach's alpha was increased to .89. The item removed was 'This app is only out for financial gain'. It was decided to remove this item not only because of reliability issues, but also because essentially every app is out for commercial gain. An example remaining item for this construct was 'I think this app protects my privacy'.

Data Sharing Attitude. The construct data sharing consisted of 4 items measuring participant's attitude towards data sharing and personalization. The Cronbach's alpha for this construct was .68, which is an acceptable reliability. An example item for this construct was 'I am concerned about personal information being sold to external parties'.

Company Trustworthiness. The construct company trustworthiness consisted of 3 items measuring how trustworthy participants felt the company behind the app was. Trust plays an important role in consumer's willingness to share personal information, thus company trustworthiness impacts consumer attitude about a company. The Cronbach's alpha measured for this construct was .75, which indicates a good reliability. An example item for this construct was ‘The company behind the app is trustworthy’.

Company Likeability. The construct company likeability consisted of 4 items measuring how likeable participants felt the company behind the app was. The likeability of a company is also an important factor in determining consumer attitude about a company (Kardes et al., 2011). The Cronbach’s alpha for this construct was .77, which indicates a a good reliability. An example item for this construct was ‘The company behind the app is concerned about people’s health’.

Intention to Download the App. The dependent variable intention to download the app consisted of one closed question, asking participants whether they wanted to download the app now. This variable was implemented to see whether there was a significant difference in participant’s intention to download the app, depending on the experimental group.

General Big Data Knowledge. The construct general big data knowledge was implemented to test how much previous knowledge participants have about big data and technology trends. It consisted of 4 items and had a Cronbach’s alpha of .68. This construct will be used as covariate in the statistical analysis. An example item was ‘I keep up to date with new digital technologies’.

Privacy Concern. The construct privacy concern consisted of only 2 items asking participants whether they read privacy statements or app permissions. This construct did not measure consumer attitude, it was implemented to gain additional information about participant’s behavior online. Thus, privacy concern will also be used as covariate in the statistical analysis. The alpha for this construct was .68, which is an acceptable reliability for a 2-item construct. An example item for this construct was ‘I read privacy statements’.

3.5 Participants

The targeted sample size for this study was 150 participants, this would guarantee that each condition had at least 20 participants. The sample is a non-probability sample based on convenience sampling as there are no specific pre-defined characteristics for the respondents. The participants were addressed via face-to-face, personal email, Facebook and survey networks such as SurveyCircle and PollPool. Each participant was assigned randomly to one of the six experimental conditions. All respondents had to be above the age of 18. The sample was composed of a total of 151 participants (64 males and 87 females. M_{age}= 29,34, ranging from 18 to 60). The study used a 2 (watching eyes: eyes vs. no eyes) x 3 (personalization: no, medium, high) between-subjects design, consisting of six different experimental conditions, as shown in Table 1 on the next page.

Table 1. Descriptives of the Participants (N=151)

Sample Characteristics (age & gender) of the six experimental conditions

		Gender		Age	
	N	Male	Female	Mean	SD
Eyes					
No Personalization	24	11	13	30,50	12,13
Medium Personalization	27	12	15	29,48	9,69
High Personalization	27	12	15	30,15	11,09
No Eyes					
No Personalization	27	13	14	30,41	11,47
Medium Personalization	24	9	15	26,17	6,95
High Personalization	22	7	15	29,05	8,50
Total	151	64	87	29,34	10,13

3.6 Test of homogeneity

To ensure that certain characteristics such as age, gender and general app evaluation indicate a homogenous distribution among the experimental conditions, tests of homogeneity using chi-square and one-way ANOVA were conducted. For the categorical variable gender a chi-square test was performed and no significant difference between the distribution of males and females at $p<0.5$ level among the experimental conditions was found, [X^2 (5, N=151) = 1.82, p=.87]. To test the homogeneity of the variable age, a one-way ANOVA was conducted and again no significant difference between the experimental groups was found, [F (5,145) =1.77, p=.12]. Furthermore, a one-way ANOVA was conducted on general app evaluation, to control for app design or usability influencing the results. Again, no significant difference was found between the experimental groups in terms of general app evaluation, [F (5, 145) =.058, p=.998]. Based on the presented results, it can be concluded that the experimental groups are homogenous regarding gender, age and general app evaluation, thus it is not necessary to perform any statistical control for these variables in further analysis.

Table 2 Comparison of the sample characteristic (age) and control variable (general app evaluation) in the six experimental conditions.

Distribution of Sample Characteristics and Control Variables

	No Personalization		Medium Personalization		High Personalization	
	Mean	SD	Mean	SD	Mean	SD
Eyes						
Age [a)]	30.50	12.13	29.48	9.69	30.15	11.09
General App Evaluation [b)]	2.50	.64	2.55	.69	2.53	.70
No Eyes						
Age [a)]	30.41	11.47	26.17	6.95	29.05	8.50
General App Evaluation [b)]	2.55	.87	2.60	.79	2.51	.75

a) Self-reported on 1 to 100 scale

b) 5-point likert scale (1=strongly agree/ 5=strongly disagree)

4 Results

4.1 App Data Safety

A three-way ANCOVA (3x2x2) was conducted with level of personalization and watching eyes as independent variables and app data safety as dependent variable. Personal health interest was included as a moderator variable. There was a significant main effect of level of personalization on app data safety [F (2,137) = 13.78, p < .01] with the results indicating that participants in the high personalization condition felt the app was less safe (M = 3.56, SD = 1.06) than participants in the no personalization condition (M = 2.66, SD = .84), as expected in H1. A post hoc analysis using LSD revealed that statistically significant differences at p<.05 level occur between the no personalization and medium personalization group (M = 3.26, SD = .86), p <.01, as well as between the no personalization and high personalization group, p <.01.

Furthermore, there was a significant main effect of personal health interest [F (1,137) = 20.81, p < .01] on app data safety. Participants with a low personal health interest felt that the app was safer (M = 2.81, SD = .94) than users with a high personal health interest (M = 3.43, SD = .95). No significant effect of watching eyes on app data safety was found [F (1,137) = .49, p=.48].

No significant interaction effects on app data safety were found. The interaction of level of personalization and watching eyes was not significant [F (2,137) = .68, p = .51], contrary to what was expected in H6. Furthermore, contrary to what was expected in H4 no significant interaction effect between level of personalization and personal health interest [F (2,137) = .13, p=.88], as well as personal health interest and watching eyes [F (1,137) = .29, p= .60] was found. An interaction of all three independent factors also showed no significant effects [F (2,137) = .82, p= .44].

4.2 Data Sharing Attitude

To analyze the effects of the independent variables level of personalization and watching eyes on the dependent variable data sharing attitude a three-way ANCOVA was conducted. Personal health interest was included as a moderator. The results indicate that there was no significant main effect of level of personalization on data sharing attitude [F (2,137) = .02, p= .98)], contrary to what was expected in H1. Also, no significant main effect of watching eyes [F (1,137) = .00, = .96] or personal health interest on data sharing attitude [F (1,137) = .07, p= .79] was found.

No interaction effects were found between level of personalization and watching eyes [F (2,137) = .58, p = .56] or between watching eyes and personal health interest [F (1, 137) = .05, p= .825]. Again, this is contrary to what was expected in H6, therefore H6 can be partly rejected regarding assumptions made in H1. Also, no interaction was found between all three independent variables [F (2, 137) = .07, p= .94]. However, there was a significant interaction effect between level of personalization and personal health interest on data sharing attitude [F (2, 137) = 6.01, p < .01]. Therefore, H4 can be partly accepted as personal health interest moderates the effect of level of personalization on data sharing attitude.

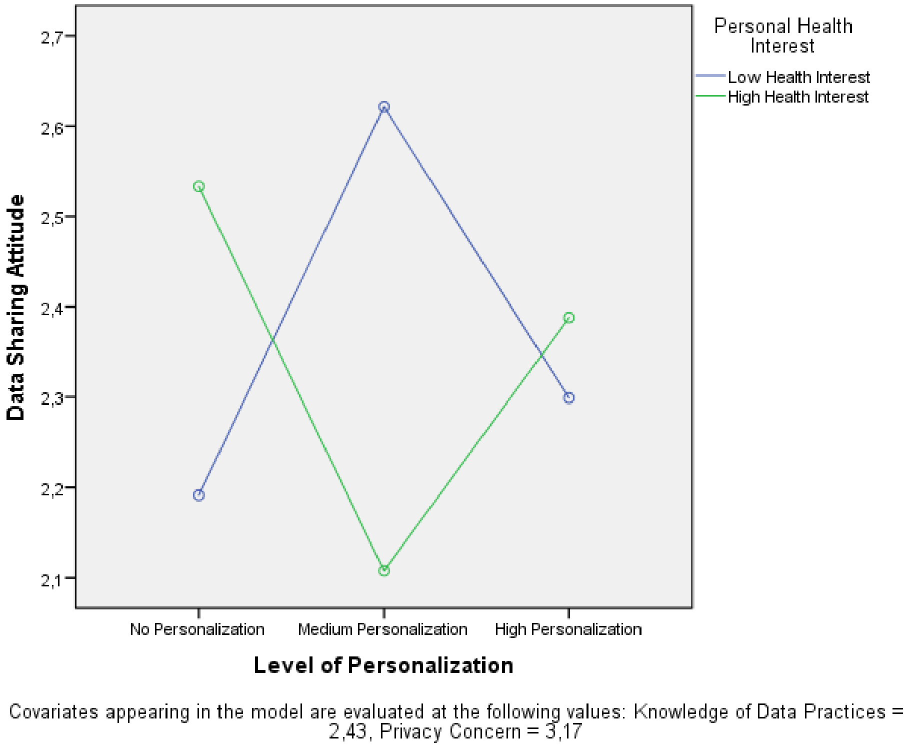


Figure 5 Data sharing attitude means by level of personalization and personal health interest.

Figure 5 shows an overview of this interaction. In the no personalization condition participants indicating a low health interest show a more concern about data sharing (M = 2.10, SD = .57) than participants with high health interest (M = 2.60, SD = .77). A similar effect can be seen in the high personalization condition, participants with low health interest show slightly more concern about data sharing (M = 2.35, SD = .51) than participants with high health interest (M = 2.42, SD = .80). In the medium personalization condition, on the other hand, participants with low health interest demonstrate much less concern about data sharing (M = 2.51, SD = .67) than participants with high health interest (M = 2.13, SD = .66).

4.3 Company Trustworthiness

Regarding the effects of company trustworthiness, a three-way ANCOVA was conducted with level of personalization and watching eyes as independent variables and company trustworthiness as dependent variable. Personal Health interest was implemented as a moderator. The effect of level of personalization on company trustworthiness was significant [F (2, 137) = 13.52, p < .01], showing that participants in the no personalization condition had a more positive attitude towards the company in terms of trustworthiness (M = 2.73, SD = .63) than participants in the high personalization condition (M = 3.47, SD = .85), as expected in H2. Again, a post hoc analysis using LSD was conducted to show that statistically significant differences at p<.05 level occur between the no personalization and medium personalization group (M = 3.22, SD = .62), p < .01, as well as the no personalization and high personalization group, p < .01.

The mean scores for company trustworthiness indicate that participants in the eyes condition (M = 3.21, SD = .81) felt the company was less trustworthy than participants in the no eyes condition (M = 3.01, SD = .73). The effect of watching eyes on company trustworthiness was marginally significant [F (1,149) =2.531, p=.114]. Moreover, a significant effect of personal health interest on company trustworthiness was found [F (1,137) = 23.92, p < .01]. Participants with high personal health interest feel that the company is less trustworthy (M = 3.34, SD = .75) than participants with low personal health interest (M = 2.83, SD = .71).

There were no interaction effects between level of personalization and watching eyes [F (2, 137) = .19, p= .83], nor between level of personalization and personal health interest [F (2, 137) = .24, p= .46]. Also, no interaction effect between watching eyes and personal health interest was found [F (1, 137) = .02, p= .89]. An interaction of all three independent variables showed no significance either [F (2,137) = .79, p= .46]. This is contrary to what was expected in H5 and H6. Finally, effect of the covariate privacy concern was significant [F (1,137) = 4.02, p= .05].

4.4 Company Likeability

A three-way ANCOVA was conducted with level of personalization and watching eyes as independent variables and company likeability as dependent variable. Personal health interest was included as moderator. The effect of level of personalization on company likeability was marginally significant [F (2,137) = 3.01, p= .06], with the mean scores indicating that participants in the no personalization condition rated the company more favorably (M= 2.75, SD = .64) than participants in the medium personalization (M= 2.91, SD= .52) or high personalization condition (M = 3.05, SD = .98), as expected in H2. A post hoc analysis using LSD was conducted to show that the difference between the no personalization and high personalization condition was indeed significant, p = .032.

Furthermore, there was a significant effect of personal health interest on company likeability [F (1,137) = 35.51, p < .01], with participants indicating a low health interest rating the company more favorably (M= 2.58, SD= .56) than participants indicating a high health interest (M= 3.16, SD= .70). However, no significant effect of watching eyes on company likeability was found [F (1,149) =.19, p=.66].

There were no interactions between level of personalization and watching eyes [F (2,137) = .52, p= .60], as well as between watching eyes and personal health interest [F (1,137) = .05, p= .83], contrary to what was expected in H6. Therefore, H6 can be rejected. However, there was a marginally significant interaction between level of personalization and personal health interest [F (2,137) = 2.12, p= .12].

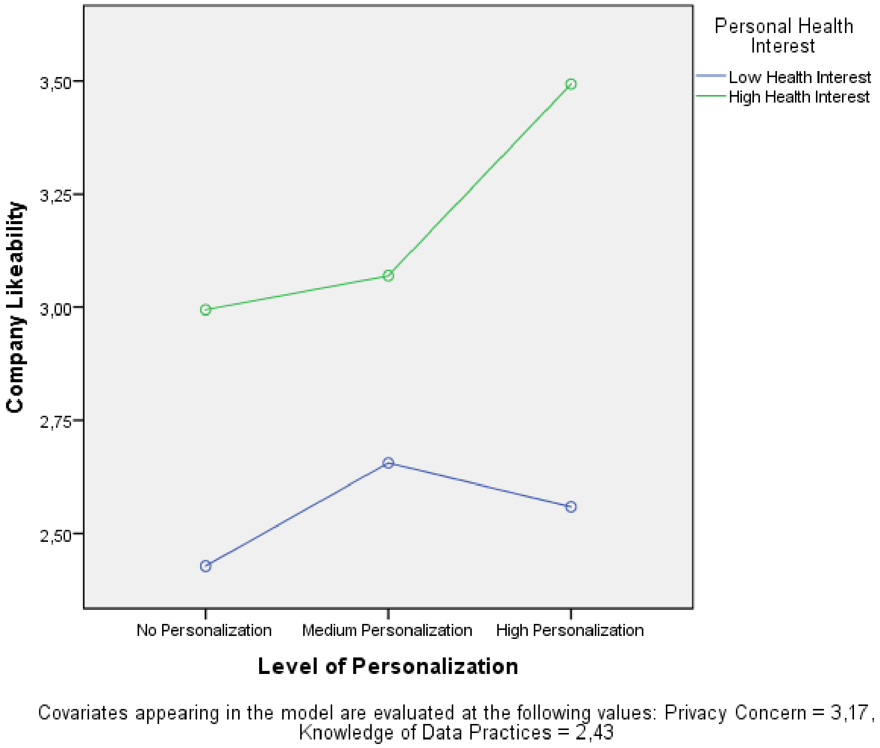


Figure 6 Company likeability means by level of personalization and personal health interest

Figure 6 shows an overview of the interaction. The results indicate that participants with high health interest generally rated the company less favorably than participants with low health interest. The largest effect can be seen in the high personalization condition, where participants with low health interest rated the company much more positively on likeability (M= 2.55, SD= .68) than participants with high health interest (M= 3.47, SD= .83). This shows that participants with higher possible interest in the app and its functions are affected more significantly by threats to their privacy. However, participants in the no personalization condition with low health interest also rated the company more favorably (M= 2.49, SD= .58) than participants with high health interest (M= 2.96, SD= .62). The same goes for the medium personalization condition, where participants with high health interest gave a more negative rating (M= 3.06, SD= .55) than participants with low health interest (M= 2.7, SD= .30). Therefore, it can be stated that personal health interest moderates the effects of level of personalization on company likeability, as expected in H5.

4.5 Intention to Download the App

To evaluate whether there was a significant difference between the experimental groups regarding participants' intention to download the app a chi-square test was conducted. The results show that there was no significant effect of the level of personalization on the intention to download the app [X^2 (2, N=151) =.05, p=.98]. Furthermore, the majority of participants (N=111, 73%) voted against downloading the app, therefore H3 can be rejected.

4.6 Qualitative Analysis of the Open Questions

In this section, the results from the open questions regarding participant's thoughts and concerns about data-driven personalization are analyzed. In order to analyze the responses to the question: *What are your thoughts about the increasing personalization of your online experience based on your personal information?* A coding scheme was developed and based on the content of the responses three coding categories were formed: *Negative thoughts about data-driven personalization, Mixed thoughts about data-driven personalization* and *Positive thoughts about data-driven personalization.* The number of participants who answered the open question was N=137. There was no significant difference between the experimental groups in the length of their responses.

The findings show that the majority of participants either think negatively (45%) about the increasing personalization of their online experience or have mixed feelings (38%). Only about 18% of the respondents think positively about increasing personalization. There were no significant differences between the experimental conditions. Generally, most participants feel a loss of control over their personal data and are highly concerned about data privacy. However, at the same time participants also acknowledge the benefits that data-driven personalization can bring. For example, this can be illustrated by the following comment: *I feel personalization of my online experience based on personal information is beneficial, but if there is a breach in what personal information is shared and the degree to which it is exposed, then it becomes an issue.* Furthermore, as a part of gaining more control over personal data participants would be willing to sell their data to companies themselves as this comment demonstrates: *Personalization is good for the user, but companies will own and sell my data. I would like to be able to sell my own data.* These findings give insight into consumer thoughts and can be used to find possible future solutions on how to handle consumer data.

In order to analyze the response to the question: *What is your main concern with regards to data sharing and personalization?* a coding scheme was developed based on the content of the responses and five coding categories were formed: *Manipulation, Criminal Activities, Misuse of Data, Loss of Privacy and No transparency.* A description of the categories is given below. There were no significant differences between the experimental groups neither in the type and occurrence of named concerns, nor the length of the responses. The number of participants who answered the open question was N=137.

The majority of participants (37%) are most concerned about the misuse of their data by a company, which mainly regards the sharing or selling of personal data to unknown third parties. In general, consumers are concerned about their data being used by companies beyond the original purpose, with consumers having no control over that usage. For example, this can illustrated by the following comment: *The misuse of my information is my main concern, like the selling of my data without informing me what data is being sold or to whom.* Another major concern by participants (23%) is the loss of privacy, as personalization requires the constant tracking and analyzing of consumer online behavior. Consumers are worried that companies can not only track them wherever they go, but also know everything about them: *My main concern is that it is possible to track me wherever I go as well as my personal likes, dislikes and status of health.* Furthermore, roughly 16% of participants are concerned about their data being used for criminal activities such as identity theft, blackmail, hacking or heavy spamming. This can by illustrated by the following comments: *Piracy, fishing or identity theft in the worst case*, or, *When data gets stolen and published by hackers.* About 12% of participants were mainly concerned about companies being able to manipulate them, based on personal consumer data, to buy certain products or services. *That someone will predict my actions, before I even start to plan them,* or, *That it is not used for commercial use anymore, but to manipulate you in more severe areas.* Finally, participants are concerned about the lack of transparency, in that they do not know what happens to their data and how it is used by companies. This can be illustrated by: *The lack of transparency concerning the storage and use of my data.*

5 Discussion

5.1 Discussion of Results

The main concern of the paper was to explore whether the facilitation of consumer understanding and awareness of data-driven personalization affects consumer attitude towards data-driven personalization and companies that engage in it. In that sense, an experiment was conducted with level of personalization and watching eyes as independent variables. The variables that were explored in terms of consumer attitude include app data safety, data sharing attitude, company trustworthiness, company likeability and intention to download the app. In this section, the results and theoretical implications of the experiment will be discussed.

This study has shown, in accordance with predictions, that an increasing level of personalization leads to a more negative attitude towards data-driven personalization in terms of app data safety. Thus, if consumers are presented with comprehensible information about a company’s data collection and sharing practices, they can recognize threats to their privacy more easily (Luzak, 2014). Hence, consumer understanding of data-driven personalization can be enhanced through the provision of clear and understandable privacy statements. Then, consumers will be able to properly judge company’s privacy practices and make informed decisions whether to accept these privacy terms or not. Furthermore, participants generally demonstrated a negative attitude towards data sharing, indicating that they are concerned about data collection and sharing practices, and try to be careful when sharing information online. These results are consistent with other studies which have shown that consumers are concerned about data-driven personalization (Quint & Rogers, 2015; Strong, 2013). Hence, the study shows that enhancing consumer understanding of a company’s data collection and sharing practices negatively affects consumer attitude towards data-driven personalization. From the consumer perspective, there is an immediate need to change how companies handle their communication about data practices, as consumers are increasingly concerned about personal data usage and demand more information.

Another interesting finding is that, contrary to the assumptions, the effect of personalization on app data safety was not moderated by personal health interest. This result was unexpected, as previous studies have shown that participants often ignore threats to privacy, if the product or service is of interest to them (Marotta-Wurgler, 2005). However, one possible reason may be that the level of safety the app provides, depends solely on the app provider and not on consumer behavior or judgment. Therefore, it is something that the consumer cannot change, as the information is controlled and provided by the company. Thus, the feeling of app data safety cannot be influenced by personal health interest. This means that personal health interest only influences variables that are directly concerned with consumer behavior or judgment, and not with factual information.

On the other hand, the study finds that participants with low health interest experiencing no personalization or high personalization show more concern about data sharing than participants indicating a high health interest. These findings are consistent with a study by Marotta-Wurgler (2005) which shows that consumers ignore possible privacy threats when a product or service interests them. In contrast, participants with low health interest experiencing medium personalization show less concern about data sharing than participants with a high health interest. This result was quite unexpected and suggests that, in contrast to the study by Marotta-Wurgler (2005), consumers do not neccesarily ignore privacy threats when a product or service interests them. One possible reason may be that the medium personalization scenario was most realistic for consumers. Indeed, the results of the pretest show that this scenario was rated as most realistic compared to the other scenarios. Thus, the invasion of privacy may have been more shocking for consumers with high health interest in the medium personalization condition, than in the other less realistic personalization conditions. However, further research exploring the effect of realism or familiarity of personalization scenarios is needed, to reach a definite conclusion.

Furthermore, the study demonstrates that an increasing level of personalization leads to a more negative attitude towards the company using data-driven personalization in terms of company trustworthiness. Thus, company trustworthiness decreases with increasing personalization. This is an important finding for companies, as brand trust plays a significant role in influencing consumer’s willingness to share personal information with a company (Gigya, 2015). Moreover, to trust a company consumers need reassurance of how personal information is used (Quint & Rogers, 2015). This means that company trustworthiness is an important component of consumer attitude towards data-driven personalization. If a consumer can trust a company to handle his or her personal information with confidentiality and only use it for the intended purpose, then he or she will be more likely to share information with that company. Thus, companies that want to be successful in the age of data-driven personalization need to invest in building consumer trust in the long run.

Furthermore, the study finds that, in accordance with predictions, increasing consumer understanding of data-driven personalization also leads to a more negative attitude in terms of company likeability. The results show that participants experiencing no personalization rated the company more favorably than participants experiencing higher levels of personalization. Again, these results agree with other studies which have shown that consumers have a negative attitude towards companies using data-driven personalization (Barnard, 2014; Strong, 2013). One reason may be that even if a company directly declares its data practices, consumers still do not have any control over the use of their personal data (Malhotra et al., 2004). This lack of control results in a more negative attitude towards the company. As of today, there are no options for consumers to completely opt-out of personal data sharing or at least have direct control over which data is shared and how it is used. This leads to a discrepancy between what companies provide and what consumers need, creating a gap. In the long run, this gap will cause consumers to dislike the company and look for alternatives elsewhere.

This study also provides evidence that, in accordance with predictions, personal health interest moderates the effect of level of personalization on company likeability. Participants with a high health interest generally rated the company less favorably on company likeability than participants with a low health interest. Especially participants with a high health interest experiencing high levels of personalization demonstrated a more negative attitude towards the company than participants with low health interest. This suggests that participants with a higher possible interest in the app are more significantly affected by threats to their privacy. The reason for that may be that this creates a discrepancy between consumer's interests, namely using the health app, and consumer's desire to exert control over his or her personal data. This discrepancy creates a negative feeling for consumers, which in turn they transfer to the company, forming a negative attitude (Kardes et al., 2011). However, as there was a general negative effect of high health interest on company likeability, further research is needed to reach a definite conclusion. Possible limitations of the research in this regard are presented later in this section.

Another interesting finding is that, in contrast to the assumptions, personal health interest does not moderate the effect of level of personalization on company trustworthiness. In general participants with low health interest perceived the company to be more trustworthy than participants with a high health interest. This suggests that consumers who are interested in the app take more factors into account for their evaluation than consumers who are not interested. Possible factors may be the design of the app or functionalities. Furthermore, in this study participants were not able to test a real app, therefore participants with a high interest in health apps could not play around with the app functions. As a result, this may have created skepticism in consumer minds about the functionality of the app, as they were only given information about the functions, without the actual proof of a testing situation. Hence, resulting in lower trust towards the company. To further explore the effect of personal health interest on company trustworthiness, additional research using a real functioning app needs to be conducted.

The implementation of watching eyes to facilitate consumer awareness did not show any significant effects. Contrary to the predictions, there was no increase in the effects of level of personalization on app data safety, data sharing attitude, company trustworthiness and company likeability due to the presence of watching eyes. However, the watching eyes had a marginally significant effect on company trustworthiness with participants in the eye condition demonstrating a more negative attitude towards the company than participants in the no eyes condition. This suggests that the watching eyes indeed increased participants awareness, indicated by a lower trust in the company, but the effects were not strong enough to impact other variables. This is contrary to the findings in other studies in which watching eyes induced a sense of being seen and caused individuals to act more pro-socially. However, these studies were conducted in a public environment and were related to social behaviors (Bateson et al., 2013; Nettle et al., 2012). The present study, on the other hand, took place in a more private environment, such as participant's private devices. This suggests that the watching eyes effect does not significantly impact human behavior in the private domain. Another reason may be that the watching eyes effect is moderated by individuals degree of public self-awareness, thereby only affecting individuals who are particularly concerned about how others see them and their social reputation (Pfattheicher & Keller, 2015). Further research is necessary to explore how consumer awareness of privacy issues can be facilitated through visual cues, to increase consumer reading of privacy statements or app permissions.

This study also shows that the level of personalization may affect consumer's willingness to share personal data when given an incentive. Participants experiencing no personalization were less likely to share personal data for a voucher, than participants experiencing higher levels of personalization. One possible reason could be that participants experiencing medium or high personalization were already primed with the concept of sharing their data with the company. Therefore, the perceived threat to their privacy was lower than for participants experiencing no personalization. However, the willingness to share personal data was not included as a valid dependent variable in this study, therefore further research is needed to support these assumptions.

Finally, in contrast to the predictions, there were no significant effects of either level of personalization or watching eyes on participant's intention to download the app. Generally, the majority of participants (73%) did not want to download the app, regardless of the experimental condition. This may relate to the possible limitation that instead of a real app, only images of an app were used for this study. A real app would have created a more realistic testing situation, as it allows participants to interact and try out the functions, increasing the reliability of the results. In this case, a long-term study using a/b testing in a real-life setting could be an interesting approach. For example, when launching a new app there are different versions of the app portrayed to the users including a comprehensible data usage declaration, normal app permissions and app permissions enabling consumers to opt-out.

Overall, this study shows that consumer understanding of data collection and sharing practices can be increased through the provision of clear and comprehensible declarations of data usage. This enables consumers to recognize threats to their privacy. Furthermore, consumers acknowledge the benefits that data-driven personalization can bring, but are significantly concerned about their privacy and losing control over personal data. In that sense, consumers' biggest concerns are the misuse of personal data beyond the originally intended purpose such as the selling or sharing of data, the loss of privacy through constant online tracking and criminal activities such as identity theft or hacking. Moreover, even if consumers are informed about threats to their privacy, they still cannot exert any control over their personal data, leading to a negative attitude towards data-driven personalization and companies engaging in it. Therefore, companies need to reconsider not only their information policy about data collection and sharing, but also think about how to give consumers more control over the usage of their personal data. But what exactly does that mean for companies?

5.2 Practical Implications

The theoretical implications discussed in the previous section lead to practical consequences. Studies have shown that consumers are becoming increasingly concerned about how their personal data is collected and shared by companies (Quint & Rogers, 2015). Until now, most companies hide their data collection and sharing practices behind extensive incomprehensible privacy statements, that are impossible for consumers to understand. However, data is on the rise of becoming today's currency with an increasing number of consumers demanding more control over the collection and sharing of personal data. Consumers are willing to share personal data in exchange for an improved and personalized online experience, as long as consumers have control over their personal information. Furthermore, brand trust is an important factor in determining consumers' willingness to share personal information. Thus, companies should become proactive by providing consumers with comprehensible and shortened privacy statements that give consumers the possibility to opt-out or to give informed consent regarding which information is collected and how it is used. An honest and straightforward information policy will build consumer trust. Additionally, to increase consumer willingness to share personal information companies could offer incentives such as vouchers or monetary rewards (Gigya, 2015).

Furthermore, consumers who are interested in a specific service or product take more factors into account when forming an attitude such as design, usability and functionality. Thus, as consumer interest in a product increases, their need for information regarding that product increases as well. Hereby, information regarding data collection and sharing also plays an important role, for the consumer to recognize privacy threats. Companies should openly present the consumer with as much information as possible, not only restricted to the product itself, but also the permissions required for usage. This enables consumers to make informed decisions and increases consumer trust in the company. Trusting consumers are more willing to share personal information.

Moreover, companies should prepare for new data collection and sharing models, such as paying consumers to collect and use their data. Data is becoming the currency of today's interconnected world, from which not only companies should benefit. With an increasing consumer understanding of data-driven personalization, consumers will soon be able to recognize the value of personal data. Companies should be prepared for that.

5.3 Limitations

The main limitation of this research was regarding the manipulated stimulus. Although, the app was designed to be as realistic as possible, it was designed by the researcher and did not represent a high quality app design. Therefore, some of the reactions by the participants may be influenced by the design of the app. More importantly, however, participants were only presented with images of the app instead of an actual real app. The researcher was only able to show participants the possible app functions, but participants were not able to interact with the app. This limitation could have possibly influenced the participants attitude towards the dependent variables, especially for participants with a high health interest. In future research a real functioning app should be used as stimulus to increase participants emotional involvement and to avoid these limitations.

Furthermore, the sample was based on convenience sampling and self-selection, therefore it cannot be guaranteed that the findings in this study are representative for a larger population. Also, when considering that the study was based on 3x2x2 factorial design due to the median split of the moderator personal health interest, a higher number of participants could have emphasized the results. Therefore, larger-scale succeeding studies that are representative of the population would be needed to confirm the results.

6 Conclusion

The aim of this study was to explore whether the facilitation of consumer awareness and understanding of a company's exact data collection and sharing practices affects consumer attitude towards data-driven personalization and the company using data-driven personalization. In conclusion, this study has shown that increasing consumer understanding of data collection and sharing processes through the presentation of clear and comprehensible declarations indeed impacts consumer attitude about data-driven personalization and companies engaging in it. It was found that consumers experiencing high personalization demonstrated a more negative attitude towards data-driven personalization. Moreover, increasing levels of personalization also had a negative impact on consumer attitude towards the company using data-driven personalization. A qualitative analysis of consumer thoughts and concerns about personalization revealed that most consumers had a more negative or mixed attitude towards data-driven personalization. Hereby, consumers' main concerns are the misuse of personal data, constant tracking as well as the criminal use of personal data.

On the other hand, the facilitation of consumer reading through visual eye cues did not prove to emphasize the effect the level of personalization. Thus, based on the results it can be concluded that consumer understanding of privacy statements is mainly facilitated by the clear and comprehensible presentation of the information.

References

Aguirre, E., Mahr, D., Grewal, D., Ruyter, K. de, & 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. doi:10.1016/j.jretai.2014.09.005

Al-Khouri, A. M. (2013). Privacy in the Age of Big Data: Exploring the Role of Modern Identity Management Systems. *World Journal of Social Science, 1*(1). doi:10.5430/wjss.v1n1p37

Barnard, L. (2014). The cost of creepiness: how online behavioral advertising affects consumer purchase intention. Retrieved from https://cdr.lib.unc.edu/indexablecontent/uuid:ceb8622f-1490-4078-ae41-4dc57f24e08b

Bateson, M., Callow, L., Holmes, J. R., Redmond Roche, M. L., & Nettle, D. (2013). Do images of 'watching eyes' induce behaviour that is more pro-social or more normative? A field experiment on littering. *PloS one, 8*(12), e82055. doi:10.1371/journal.pone.0082055

Bertolucci, J. (2013). Big Data Analytics: Descriptive Vs. Predictive Vs. Prescriptive: What distinguishes these three key types of analytics? A data scientist explains the differences. Retrieved from http://www.informationweek.com/big-data/big-data-analytics/big-data-analytics-descriptive-vs-predictive-vs-prescriptive/d/d-id/1113279

Chia, P. H., Yamamoto, Y., & Asokan, N. (2012). Is this app safe? In A. Mille, F. Gandon, J. Misselis, M. Rabinovich, & S. Staab (Eds.): *ACM Digital Library, Proceedings of the 21st international conference on World Wide Web* (p. 311). New York, NY: ACM.

Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., & Vlaev, I. (2012). Influencing behaviour: The mindspace way. *Journal of Economic Psychology, 33*(1), 264–277. doi:10.1016/j.joep.2011.10.009

eMarketer.com. (2013). Big Data Helps Reveal Consumer Behavior: Marketers, agencies on same page regarding benefits from data analysis. Retrieved from http://www.emarketer.com/Article/Big-Data-Helps-Reveal-Consumer-Behavior/1010357

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research, 69*(2), 897–904. doi:10.1016/j.jbusres.2015.07.001

EY. (2013). *The Big Data Backlash*. Retrieved from http://www.ey.com/Publication/vwLUAssets/EY-The-Big-Data-Backlash/$FILE/EY-The-Big-Data-Backlash.pdf

Gigya. (2015). The 2015 State of Consumer Privacy & Personalization.

Haley, K. J., & Fessler, D. M. (2005). Nobody's watching? *Evolution and Human Behavior, 26*(3), 245–256. doi:10.1016/j.evolhumbehav.2005.01.002

Hill, S. (2015). How much do online advertisers really know about you? We asked an expert. Retrieved from http://www.digitaltrends.com/computing/how-do-advertisers-track-you-online-we-found-out/#ixzz4Lkog72jP

Jensen, C., Potts, C., & Jensen, C. (2005). Privacy practices of Internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies, 63*(1-2), 203–227. doi:10.1016/j.ijhcs.2005.04.019

Kardes, F. R., Cline, T. W., & Cronley, M. L. (2011). Consumer behavior: Science and practice (International ed.). [Mason, Ohio]: South-Western Cengage Learning.

Labrecque, L. I., vor dem Esche, J., Mathwick, C., Novak, T. P., & Hofacker, C. F. (2013). Consumer Power: Evolution in the Digital Age. *Journal of Interactive Marketing, 27*(4), 257–269. doi:10.1016/j.intmar.2013.09.002

Luzak, J. A. (2014). Privacy Notice for Dummies?: Towards European Guidelines on How to Give “Clear and Comprehensive Information” on the Cookies’ Use in Order to Protect the Internet Users’ Right to Online Privacy. *Journal of Consumer Policy, 37*(4), 547–559. doi:10.1007/s10603-014-9263-3

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research, 15*(4), 336–355. doi:10.1287/isre.1040.0032

Marotta-Wurgler, F. (2005). *Understanding Privacy Policies: Content, Self-Regulation, and Market Forces*. Retrieved from http://www.law.uchicago.edu/files/file/marotta-wurgler_understanding_privacy_policies.pdf

McDonald, A. M., & Cranor, L. F. (2008). The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society*. Retrieved from http://lorrie.cranor.org/pubs/readingPolicyCost-authorDraft.pdf

Meinert, D. B., Peterson, D. K., Criswell, J. R., & Crossland, M. D. (2006). Privacy Policy Statements and Consumer Willingness to Provide Personal Information. *Journal of Electronic Commerce in Organizations, 4*(1), 1–17. doi:10.4018/jeco.2006010101

Meng, W., Ding, R., Chung, S. P., Han, S., & Lee, W. (2016). The Price of Free: Privacy Leakage in Personalized Mobile In-App Ads. *NDSS*. Retrieved from http://www.cc.gatech.edu/~wmeng6/ndss16_mobile_ad.pdf

Mick, D. G., & Fournier, S. (1998). Paradoxes of Technology: Consumer Cognizance, Emotions, and Coping Strategies. *Journal of Consumer Research, 25*(2), 123–143. doi:10.1086/209531

Miyazaki, A. D. (2008). Online Privacy and the Disclosure of Cookie Use: Effects on Consumer Trust and Anticipated Patronage. *Journal of Public Policy & Marketing, 27*(1), 19–33. doi:10.1509/jppm.27.1.19

Morey, T., Forbath, T., & Schoop, A. (2015). Customer Data: Designing for Transparency and Trust. Retrieved from https://hbr.org/2015/05/customer-data-designing-for-transparency-and-trust

Nettle, D., Nott, K., & Bateson, M. (2012). 'Cycle thieves, we are watching you': impact of a simple signage intervention against bicycle theft. *PloS one, 7*(12), e51738. doi:10.1371/journal.pone.0051738

O'Connor, P. (2007). Online Consumer Privacy: An Analysis of Hotel Company Behavior. *Cornell Hotel and Restaurant Administration Quarterly, 48*(2), 183–200. doi:10.1177/0010880407299541

Oracle. (2012). Big Data Analytics Technology Brief: Customer Segmentation Engines as Building Block. Retrieved from http://www.oracle.com/us/technologies/big-data/bda-customer-segmentation-engines-2045188.pdf

Orwell, G. (1949). *1984*: Seckers & Warburg.

Pariser, E. (2011). The Filter Bubble: What the Internet Is Hiding from You. London: Penguin Books.

Peterson, A. (2014). Privacy advocates warn of ‘nightmare’ scenario as tech giants consider fitness tracking. Retrieved from https://www.washingtonpost.com/news/the-switch/wp/2014/05/19/privacy-advocates-warn-of-nightmare-scenario-as-tech-giants-consider-fitness-tracking/?utm_term=.c53997dc6d7b

Pfattheicher, S., & Keller, J. (2015). The watching eyes phenomenon: The role of a sense of being seen and public self-awareness. *European Journal of Social Psychology, 45*(5), 560–566. doi:10.1002/ejsp.2122

Predictive Analytics Today. (2014). What is Predictive Analytics? Retrieved from http://www.predictiveanalyticstoday.com/what-is-predictive-analytics/

Quint, M. & Rogers, D. (2015). What is the Future of Data Sharing?: Consumer Mindset and the Power of Brands. Retrieved from https://www8.gsb.columbia.edu/globalbrands/research/future-of-data-sharing

Reachforce. (2015). How Big Data is Changing the Face of Market Segmentation. Retrieved from http://www.reachforce.com/blog/how-big-data-is-changing-the-face-of-market-segmentation/

SAS. (2016a). Big Data Analytics: What it is any why it matters. Retrieved from http://www.sas.com/en_us/insights/analytics/big-data-analytics.html#dmtoday

StatCounter. (2016). Mobile and tablet internet usage exceeds desktop for first time worldwide. Retrieved from http://gs.statcounter.com/press/mobile-and-tablet-internet-usage-exceeds-desktop-for-first-time-worldwide

Statista. (2015). Share of internet users who use health and fitness apps every month as of 3rd quarter 2015. Retrieved from https://www.statista.com/statistics/502195/health-and-fitness-app-access/

Stoicescu, C. (2015). Big Data, the perfect instrument to study today's consumer behavior. *Database Systems Journal*, *4*(3), 28–41. Retrieved from http://www.dbjournal.ro/archive/21/21_4.pdf

Strong, C. (2013). The big data arms race part two: consumer perceptions. Retrieved from https://www.theguardian.com/media-network/media-network-blog/2013/oct/04/consumer-marketing-big-data-perceptions

Teltzrow, M., & Kobsa, A. (2004). Impacts of User Privacy Preferences on Personalized Systems. In C.-M. Karat, J. O. Blom, & J. Karat (Eds.), *Human-computer interaction series: Vol. 5. Designing Personalized User Experiences in eCommerce* (pp. 315–332). Dordrecht: Springer Netherlands.

Tene, O. & Polonetsky, J. (2012). Privacy in the Age of Big Data: A Time for Big Decisions. Retrieved from http://www.stanfordlawreview.org/online/privacy-paradox/big-data

van Rijmenam, M. (2014). The Future of Big Data:: Prescriptive Analytics Changes the Game. Retrieved from http://data-informed.com/future-big-data-prescriptive-analytics-changes-game/

Zang, J., Dummit, K., Graves, J., Lisker, P., & Sweeney, L. (2015). *Who Knows What About Me? A Survey of Behind the Scenes Personal Data Sharing to Third Parties by Mobile Apps*. Retrieved from http://techscience.org/a/2015103001

Appendix A: The Questionnaire

Dear Participant,

For my master thesis in Marketing Communications at the University of Twente, I am conducting research on health apps that can help you become proactive in managing your health via your smartphone.

For the purpose of this study I designed a prototype of the health app “Superfit Me”. With this app you can easily keep track of your overall level of fitness, training success and eating habits. Today I would like to gather your feedback on the usability and design of the app. First I will show you some screenshots to give you an idea about the app functions and ask some questions about your first impression. Afterwards you will be asked to test the app prototype.

This survey will only take you about 5-7 minutes to complete. Please make sure that you read all the information provided carefully. Rest assured that all the input you provide here will be strictly confidential and anonymous.

As a token of appreciation you will have the chance to win a 20€ Amazon voucher!

Thank you in advance and I really appreciate your help!

Best,
Pascale Bastian

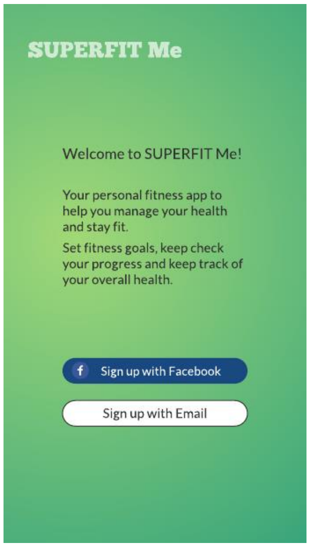
Please click "Next" to proceed with the survey.

Please read the information provided below carefully and then proceed via the "Next" button:

Imagine that you ride your bike to university or work every day, because it is convenient and it keeps you active. You are interested in leading a healthy lifestyle, therefore, decide to download a health app that lets you keep track of your physical activity and eating habits.

The app enables you to track your biking routine and set training goals. Furthermore, it uses location services to track the distance, speed and timing of your bike rides. By combining your training information with your gender, age, height and weight the app calculates your overall level of fitness and lets you monitor any changes. Additionally, you can track your eating habits by recording your meals and calorie intake.

Here you can see the prototype "Welcome" screen with your login options. Please indicate below which login option you prefer and then proceed via the "Next" button.



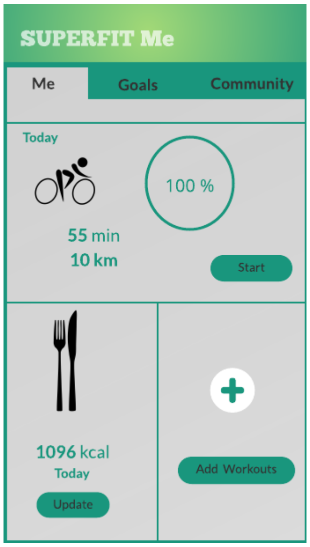
Welcome to SUPERFIT Me!

Your personal fitness app to help you manage your health and stay fit.

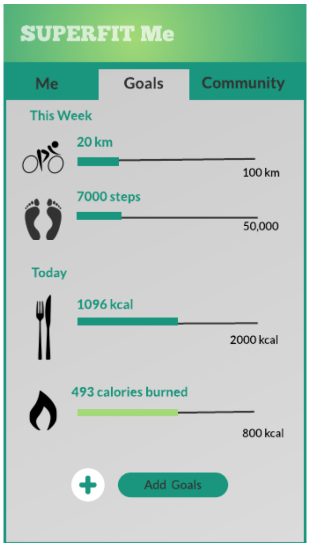
Set fitness goals, keep check your progress and keep track of your overall health.

Which login option do you prefer?
( ) Sign up with Facebook
( ) Sign up with Email
( ) Both

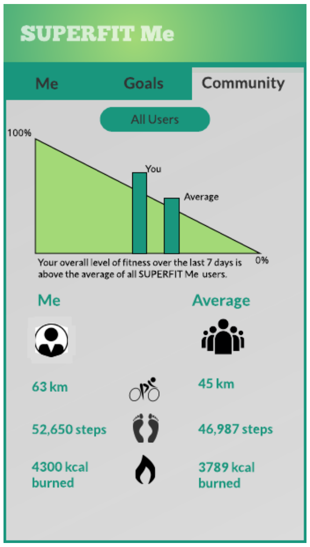
Once you are logged in you will see the "Me" screen, which gives you a great overview of your recent training sessions and eating behaviour. Here you also have the option to add more functions to the screen including additional workouts or health measures such as weight or blood pressure.



In the second tab you see the "goals" screen, which gives you a great overview of your weekly and daily goals. Here you also have the option to add more goals or to change the time frame for achieving your goals.



In the third tab you can compare your level of fitness to other members of the "SUPERFIT Me" community. Here you also have the option to compare yourself to your friends or change the time frame of the comparison (daily, weekly or monthly) by clicking on the "All Users" button.



Your overall level of fitness over the last 7 days is above the average of all SUPERFIT Me users.

Stimulus (see Appendix B)

Please indicate to what extent you agree or disagree with the following statements.

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
I am interested in health apps.	( )	( )	( )	( )	( )
I like the design of the app.	( )	( )	( )	( )	( )
This app is safe to use.	( )	( )	( )	( )	( )

This app can help me lead a healthy lifestyle.	( )	( )	( )	( )	( )
I would use this app to track my fitness.	( )	( )	( )	( )	( )
I think this app protects my privacy.	( )	( )	( )	( )	( )
The app looks easy to use.	( )	( )	( )	( )	( )
I would download this app.	( )	( )	( )	( )	( )
I feel safe providing my personal information to this app.	( )	( )	( )	( )	( )
I would be interested in finding out more about the app.	( )	( )	( )	( )	( )
This app is only out for commercial gain.	( )	( )	( )	( )	( )

Please indicate to what extent you agree or disagree with the following statements.

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
The company behind the app is concerned about people’s health.	( )	( )	( )	( )	( )
The company behind the app is trustworthy.	( )	( )	( )	( )	( )
The company behind the app is reliable.	( )	( )	( )	( )	( )
The company behind the app is only out for financial gain	( )	( )	( )	( )	( )
The company behind the app is concerned about privacy.	( )	( )	( )	( )	( )
The company behind the app is likeable.	( )	( )	( )	( )	( )
I would like to find out more about the company behind the app.	( )	( )	( )	( )	( )

Would you like to download the "SUPERFIT Me" now?

( ) Yes

( ) No

Sorry, you will not be able to download the app here, as this study is not actually about testing an app. My research is about how awareness of the extent and implications of big data personalization and data sharing affects your attitude towards these practices and companies engaging in them. So thank you for your help so far!
To round up my research, I still have a few general questions for you, so please proceed via the "Next" button.

Please indicate to what extent you agree or disagree with the following statements.

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
I keep up to date with new apps.	( )	( )	( )	( )	( )
I read privacy statements of websites I visit.	( )	( )	( )	( )	( )
I keep up to date with new digital technologies.	( )	( )	( )	( )	( )
I know what Big Data means.	( )	( )	( )	( )	( )
Selling personal information to personalize my web experience is acceptable.	( )	( )	( )	( )	( )
I am careful when sharing information online.	( )	( )	( )	( )	( )
I read app permissions carefully.	( )	( )	( )	( )	( )
I am concerned about personal information being sold to external parties.	( )	( )	( )	( )	( )
I am aware of current online privacy issues.	( )	( )	( )	( )	( )
I share a lot of personal information online.	( )	( )	( )	( )	( )

How many apps do you have on your smartphone?

( ) 1-10

( ) 11-19

( ) more than 20

( ) I don’t own a smartphone

BIG BROTHER IS WATCHING YOU

What are your thoughts about the increasing personalization of your online experience based on your personal information?

What is your main concern with regards to data sharing and personalization?

What is your gender?

( ) Male

( ) Female

How old are you?

If you would like the chance to win a 20€ Amazon voucher, please enter your email address here:

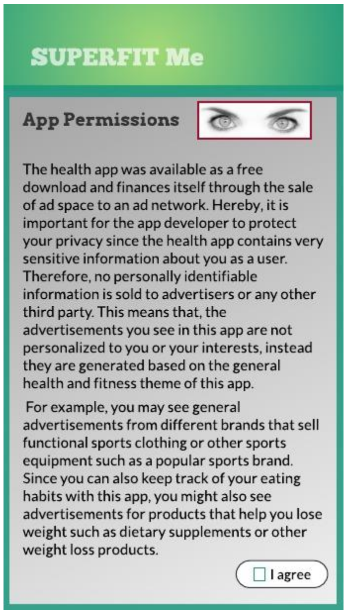
THE END

Thank you very much for participating in my survey!

If you would like to find out more about my research or have questions regarding this survey, please don't hesitate to contact me at p.bastian@student.utwente.nl .

Best,
Pascale Bastian

Appendix B : The Stimulus



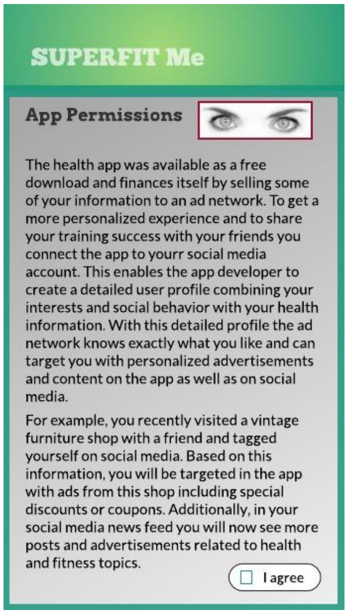
SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself through the sale of ad space to an ad network. Hereby, it is important for the app developer to protect your privacy since the health app contains very sensitive information about you as a user. Therefore, no personally identifiable information is sold to advertisers or any other third party. This means that, the advertisements you see in this app are not personalized to you or your interests, instead they are generated based on the general health and fitness theme of this app.

For example, you may see general advertisements from different brands that sell functional sports clothing or other sports equipment such as a popular sports brand. Since you can also keep track of your eating habits with this app, you might also see advertisements for products that help you lose weight such as dietary supplements or other weight loss products.

□ I agree



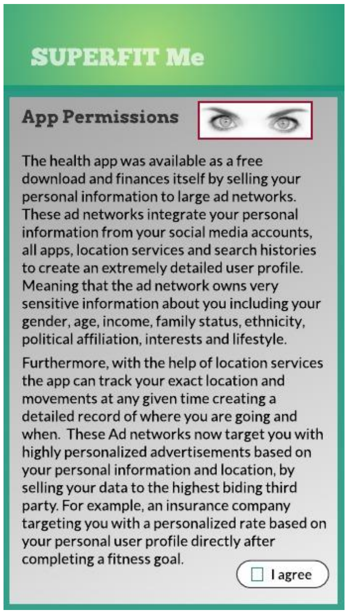
SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself by selling some of your information to an ad network. To get a more personalized experience and to share your training success with your friends you connect the app to yourr social media account. This enables the app developer to create a detailed user profile combining your interests and social behavior with your health information. With this detailed profile the ad network knows exactly what you like and can target you with personalized advertisements and content on the app as well as on social media.

For example, you recently visited a vintage furniture shop with a friend and tagged yourself on social media. Based on this information, you will be targeted in the app with ads from this shop including special discounts or coupons. Additionally, in your social media news feed you will now see more posts and advertisements related to health and fitness topics.

□ I agree



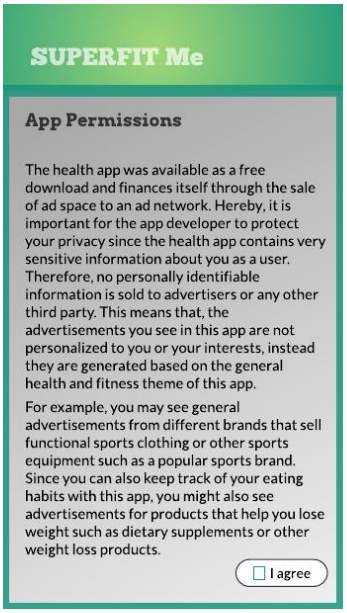
SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself by selling your personal information to large ad networks. These ad networks integrate your personal information from your social media accounts, all apps, location services and search histories to create an extremely detailed user profile. Meaning that the ad network owns very sensitive information about you including your gender, age, income, family status, ethnicity, political affiliation, interests and lifestyle.

Furthermore, with the help of location services the app can track your exact location and movements at any given time creating a detailed record of where you are going and when. These Ad networks now target you with highly personalized advertisements based on your personal information and location, by selling your data to the highest biding third party. For example, an insurance company targeting you with a personalized rate based on your personal user profile directly after completing a fitness goal.

□ I agree



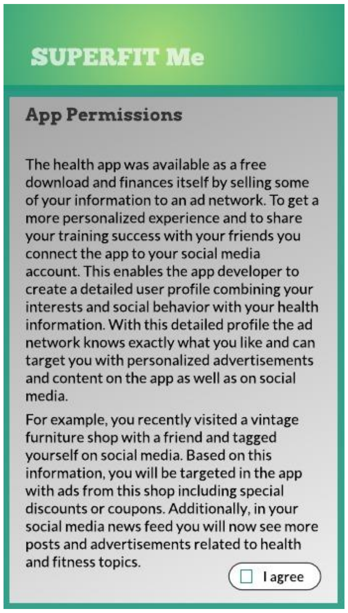
SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself through the sale of ad space to an ad network. Hereby, it is important for the app developer to protect your privacy since the health app contains very sensitive information about you as a user. Therefore, no personally identifiable information is sold to advertisers or any other third party. This means that, the advertisements you see in this app are not personalized to you or your interests, instead they are generated based on the general health and fitness theme of this app.

For example, you may see general advertisements from different brands that sell functional sports clothing or other sports equipment such as a popular sports brand. Since you can also keep track of your eating habits with this app, you might also see advertisements for products that help you lose weight such as dietary supplements or other weight loss products.

□ I agree



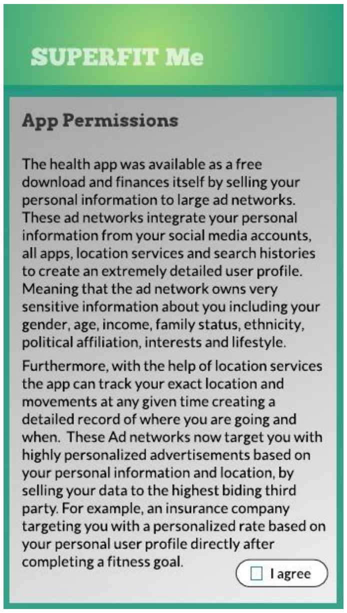
SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself by selling some of your information to an ad network. To get a more personalized experience and to share your training success with your friends you connect the app to your social media account. This enables the app developer to create a detailed user profile combining your interests and social behavior with your health information. With this detailed profile the ad network knows exactly what you like and can target you with personalized advertisements and content on the app as well as on social media.

For example, you recently visited a vintage furniture shop with a friend and tagged yourself on social media. Based on this information, you will be targeted in the app with ads from this shop including special discounts or coupons. Additionally, in your social media news feed you will now see more posts and advertisements related to health and fitness topics.

□ I agree



SUPERFIT Me

App Permissions

The health app was available as a free download and finances itself by selling your personal information to large ad networks. These ad networks integrate your personal information from your social media accounts, all apps, location services and search histories to create an extremely detailed user profile. Meaning that the ad network owns very sensitive information about you including your gender, age, income, family status, ethnicity, political affiliation, interests and lifestyle.

Furthermore, with the help of location services the app can track your exact location and movements at any given time creating a detailed record of where you are going and when. These Ad networks now target you with highly personalized advertisements based on your personal information and location, by selling your data to the highest biding third party. For example, an insurance company targeting you with a personalized rate based on your personal user profile directly after completing a fitness goal.

□ I agree

Appendix C: The Pretest

Imagine that you ride your bike to university or work every day, because it is convenient and it keeps you active. You are interested in leading a healthy lifestyle, therefore, decide to download a health app that lets you keep track of your physical activity and eating habits. With the app you can easily track your biking routine and see whether you are improving or you need to push yourself more. You can also set personal goals for training whether its reaching a new best time or weight loss. The app helps you track the distance, speed and timing of your bike ride and calculate your calorie burn. Combined with information about your weight, height, gender and age the app can analyze your overall level of fitness and any improvements over time. To accurately track the distance and save your route the app uses GPS location services. Furthermore, the app helps you keep track of your eating habits by recording all your meals. This gives you a great overview of your calorie intake and how healthy your food choices are.

Scenario 1

No personalized advertisements. Advertisements only based on the purpose of the app – sports, health and weight loss – Topic targeting ads are related to the app.

The health app was available as a free download and finances itself through the sale of ad space to an ad network. Hereby, it is important for the app developer to protect the privacy of the users since the health app contains sensitive user information. Therefore, no personal user information is sold to advertisers or any other third party. This means that, the advertisements you see in the app are not personalized to you or your interests, instead they are generated based on the general health and fitness theme of the app. For example, you may see general advertisements from different brands that sell functional sports clothing or other sports equipment such as a popular sporting brand. Since you can also keep track of your eating habits with the app, you might also see advertisements for products that help you lose weight such as dietary supplements or other weight loss products. 151 words

Scenario 2

Lightly personalized advertisements based on your interests (eg. Riding bikes instead of jogging). Interest targeting ads related to interests instead of only general topic.

The health app was available as a free download and finances itself through the sale of ad space to an ad network. The app will provide the ad network with general information about your interests such as your training preference and regional location. However, no personal or sensitive information is shared with any third parties. The ad network can now provide you with more relevant and personalized advertisements that are targeted to your needs. For example, your main training goal is to lose weight so instead of functional outdoor clothing, the app will show you personalized advertisements for weight loss or health products that help you attain your goal. For example, you live in the Netherlands in an area where it rains a lot and use your bike as main means of transportation, based on that personal data the app will show you advertisements of functional outdoor clothing especially designed for cyclists. 152 words

Scenario 3

Medium personalized advertisements based on your interests and your personal information such as gender, age and income.

The health app was available as a free download and finances itself by selling some of your information to an ad network. To get a more personalized experience and to share your training success with your friends you connect the app to your Facebook account. This enables the app developer to create a detailed user profile combining your interests and social behavior with your health information. With this detailed profile the ad network knows exactly what you like and can target you with personalized ads and content on the app and on social media. For example, you recently visited a vintage furniture shop with a friend and tagged yourself on social media. Based on this information, you will be targeted in the app ads from this shop including special discounts or coupons. Additionally, in your social media news feed you will now see more posts and advertisements related to health and fitness topics. 153 words

Scenario 4

Highly personalized advertisements based on your interests, demographic information (gender, age, income) and geographical location.

The health app was available as a free download and finances itself by selling your personal information to an ad network. This ad networks integrate your personal information from your health app, social media accounts, location services and search histories to create a detailed user profile. The ad network knows exactly how fit or healthy you are, who your friends are, what you like and where you are. This means that you can be targeted with content and ads that are customized to your personal interests, ideals and location. For example, the app tracks you on your daily bike ride to work, based on your search history you are looking for a new wardrobe. On your way to work you pass by a furniture store and in that moment, receive an ad on your app as well as on social media including a discount for wardrobes in that store. 149 words

Scenario 5

Extremely personalized advertisements based on your interests, demographic information (gender, age, income, political affiliation), geographical location and sharing of your information with third parties through ad networks.

The health app was available as a free download and finances itself by selling your personal information to large ad networks. These ad networks integrate your personal information from your social media accounts, all apps, location services and search histories to create an extremely detailed user profile. Meaning that the ad network owns very sensitive information about yourself including your gender, age, income, family status, ethnicity, political affiliation, interests and lifestyle choices. Furthermore, with the help of location services the app can also track your exact location and movements at any given time creating a detailed record of where you are going and when. Ad networks can now target you with highly personalized advertisements based on your personal information and location, by selling your information to the highest biding third party. For example, an insurance company targeting you with a personalized rate based on your personal user profile directly after completing a fitness goal. 153 words