ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
Bachelor Thesis Economics & Business Economics

The effect of confrontation with data collection on social media usage

Author: Emil E. Y. de Jong

Student number: 532158

Thesis Supervisor: Dr. S. Beltman

Second Reader: to be announced

Date: June 25th, 2024

Photo credits front page: https://www.tripwire.com/state-of-security/data-privacy-day-understanding-risks-social-media

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second reader, Erasmus School of Economics or Erasmus University Rotterdam.

Abstract:

In this thesis, I study the effect of confrontation with data collection on the intention to make a social media account. Through interviews an understanding is formed of the importance of privacy and other factors while making a social media account. Through a survey an experiment is conducted, where the first group was confronted with the data collection and the second group is not. It was found that Dutch residents of 20 -25 years have a strong negative reaction to the confrontation with data collection. It could not be determined whether this was caused by increased privacy awareness or preferences. When people are reminded of the Data that is collected of them, they will often opt-out of the data collection.

Keywords: Social Media, Intention to Use, Privacy, Experiment, Confrontation with Data Collection.

Table of Contents

List of Figures & Tables

Chapter 1: Introduction

1.1 Problem definition

This research is about awareness of data privacy, related concerns and the intention to make a social media account. The UN has stated that in the age of digitalization, privacy is becoming an increasingly more important asset that is increasingly difficult to obtain (OHCHR, 2022). The data collection by social media companies breaches this privacy. Privacy is important to retain independence in who you are and what you do (AP, n.d.). Social media companies collect: location data, even when the location services are turned off by using IP-address, data about the used device such as battery level, other installed apps and files names and types, and meta data about the use of the app and the pictures on the device it has access to (Paragraph 2.5). People do care about their privacy (Gutierrez, 2019). However, even people who indicate to care a lot about privacy do not act more privacy conscious (Acquisti & Grossklags, 2005; Barnes, 2006; Norberg, et al., 2007; Gerber et al., 2018; Willems et al. 2023; Dinev, 2024). The question arises: what will make people act on their privacy awareness and preferences when it comes to data sharing?

The relation between personal preferences and awareness of certain topics and resulting behavior has been researched in many fields. For instance, the influence of ecological awareness and ecological consumer behavior on intention to purchase (Paul & Rana, 2012). The role of data privacy in marketing is also a widely researched topic (Martin & Murphy, 2017). However, when privacy concerns are made salient, the targeting is ineffective (Goldfarb & Tucker, 2011). In the literature it is also found that there seems to be a privacy paradox. Even though, people indicate to care about their privacy, they do not act accordingly by ensuring their privacy (Acquisti & Grossklags, 2005; Barnes, 2006; Norberg, et al., 2007; Gerber et al., 2018; Willems et al. 2023; Dinev, 2024). Additionally academic research indicated the collection of data can be problematic (Janek, 2022).

Previous research into people's intentions and the effect of privacy concerns has been conducted. This was in the context of purchasing products or to use a service. Resp. (George et al., 2021) and (Mariani et al., 2021; Choi et al., 2023). The intention to use social media has also been researched, however privacy is usually not considered as a relevant factor (Yuan et al., 2021 and Balakrishnan, 2016). Privacy concerns were mentioned in one survey question in the research of Arulogun et al. (2020), but not given any notion in the rest of their research paper. The problem researched in this study will be the effect of privacy on social media use.

1.2 Research scope and question

To the best of my knowledge, the relation between intention to use social media and privacy has not been researched. It is important that this relation receives investigative attention.. For consumers it is important to know that they can act on their privacy awareness and preferences, for social media companies it is important to know under what condition people will opt out of their service for privacy reasons. For legislators this is also important in order to make effective privacy legislation. Dutch residents are widely available for the research. The choice for a fictional social media account is made to eliminate prejudice about certain social media, but it should be noted that a social medium like Instagram is intended. The age cohort of 20 to 25 was the first generation to grow up with social media widely available, they are the first generation to constantly have their data collected by social media companies Therefore, the central question that is researched in this thesis is:

How will Dutch residents from 20 – 25 years react to confrontation with data collection practices, when making a fictional social media account?

To answer this question, the following sub questions need to be answered qualitatively and quantitatively:

Sub questions for the qualitative research:

1. What do Dutch residents from various ages consider when making a social media account, such as Instagram?
2. What are possible reactions to confrontation with the data collection practices from social media companies?
3. How do privacy awareness and preferences influence the intention to make a social media account?

Sub questions for the quantitative research:

4. Are Performance, Effort/influence, Self and Communication Functionality also applicable in a general social media context as explaining factors?
5. Does a reminder on privacy increase the privacy awareness and change preferences?
6. Do Performance, Effort/influence, Self and Communication Functionality explain the intention to make a social media account?
7. Do privacy awareness and preferences explain the intention to make a social media account?

1.3 Method

For this thesis two research methods are used. First qualitative interviews are thematically analyzed. Insights in social media use and privacy are gained there. This partly builds the foundation for the privacy related questions of the second part which is a quantitative survey about the factors that decide whether people will make a social media account. Independent Sample t-testing, factor analysis and linear regression are used to evaluate the results of the survey.

1.4 Hypotheses

The main hypothesis in this study is that people are more inclined to refuse social media when they are confronted with the data that is collected from them. This effect will likely be stronger for people who are more privacy aware than their less privacy aware counterparts. The results may unravel the privacy paradox, by demonstrating that an imminent privacy reminder - highlighting the risks in their risk benefit analysis - may disable the paradoxical state that people are privacy aware but do not act on their privacy awareness and preferences. The hypothesis is developed based on the theoretical framework (Chapter 2).

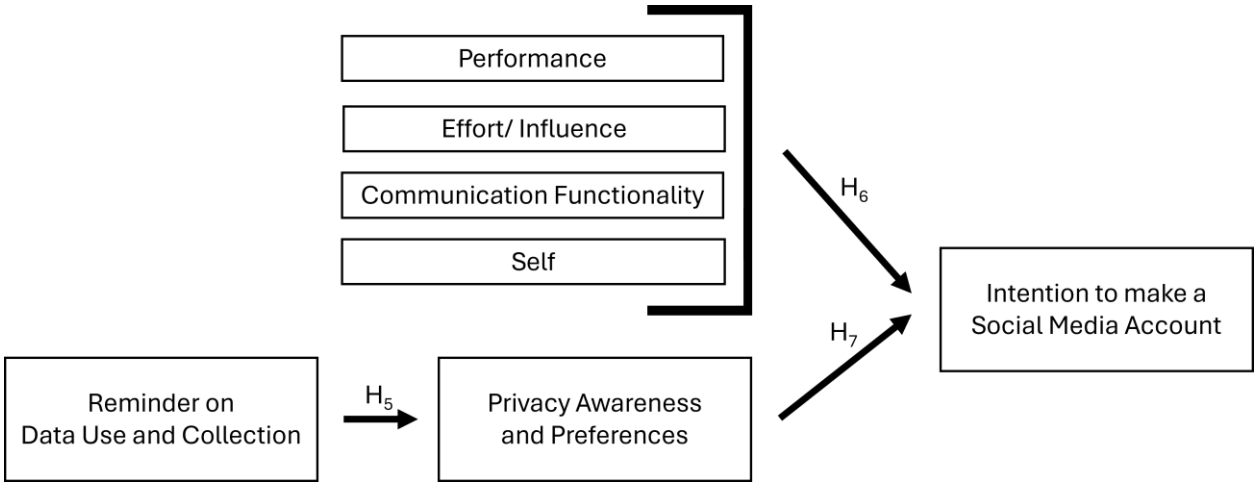
H_1: The hypothesis is that when making a social media account people will consider Performance, Effort/influence, Self and Communication Functionality (Balakrishnan, 2016).

H_2: The expectation is that people will act like they do under the theory of the privacy paradox (Acquisti & Grossklags, 2005; Barnes, 2006; Norberg, et al., 2007; Gerber et al., 2018; Dinev, 2024, p. 99) and the privacy calculus model (Laufer & Wolfe, 1977; Culnan & Armstrong, 1999). Meaning that the respondents will see issues in the data collection, but they will not completely change their behavior based on the new information they have.

H_3: The hypothesis is that privacy awareness and preferences will play a moderately important role in their decision to make a social media account. It will be considered, because people are concerned about privacy (Gutierrez, 2019). However, they will also take into account the benefits of having a social media account (Laufer & Wolfe, 1977; Culnan, & Armstrong 1999).

H_4: The hypothesis is that the factors of Balakrishnan (2016): Performance, Effort/influence, Self and Communication Functionality will be applicable in a more general context of social media use, when the questions to find the factors are properly adapted. The determinants will stay the same because the determinants describe general hurdles and benefits of social media.

Figure 1
Hypothesis overview



H_5: Following the implications of the study by Gravert (2021) a reminder about a goal can change behavior. This happens in two steps. First, if the goal is made more salient, this changes the cost benefit analysis of the respondent and then consequently this changes the behavior. It is hypothesized that a reminder on date use and collection by a social media company will increase the immediate privacy awareness and strengthen existing preferences.

H_6: Following from the Social Media Acceptance Model (SMAM) from Balakrishnan (2016) social media use for educational purposes is determined by Performance, Effort/influence, Self and Communication Functionality for 71%. The questions need to be adapted to a more general context, than just for educational purposes. The hypothesis is that these determinants are applicable to explain general social media use and that the decision to make a Social media account is heavily influenced by these four determinants.

$H_{7:}$ Privacy awareness and preferences will have a negative correlation with the intention to make a social media account. This follows from the research by George et al. (2021) and Choi et al. (2023), where privacy awareness and preferences are considered as a determinant for behavior. And from the general consumer concerns with data collection by social media companies (Gutierrez, 2019).

1.5 Thesis structure

Chapter 2 will explain the theoretical concepts used in this thesis. Chapter 3 gives a comprehensive overview of the data and the research methodology per research method. Chapter 4 presents and discusses the results of the research. Chapter 5 summarizes the research and presents the main findings, explains the implications and discusses the limitations of this research and make recommendations for future research.

Chapter 2: Theoretical framework

This chapter elaborates on the most important theoretical concepts used in this thesis. Paragraph 2.1 supports hypothesis 1, 4 and 6 and describes the basis for the main assumptions and relation studied in this thesis and provides a theoretical basis for the research question. The main concept that is used in this thesis is privacy, which is described in paragraph 2.2. This paragraph supports hypothesis 2 and 3. The understanding of the relation between privacy and making a social media account are discussed in paragraph 2.3, this supports hypothesis 7. Paragraph 2.4 discusses the confrontation with data collection, this supports hypothesis 5.

2.1 Intention of use social media

The intention to use social media lays in the acceptance of the service. Balakrishnan (2016) has proposed a Social Media Acceptance Model (SMAM) based on the Electronic Learning Acceptance Model (ELAM) (Umrani-Khan & Iyer, 2009) and Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003). These two models themselves originate in the Technology Acceptance Model (TAM) created by Davis (1989). While making the TAM, Davis (1989) found that ease of use is a determinant of perceived usefulness. Meta-analysis of the TAM underlines it robustness and powerfulness (King & He, 2006).

The UTAUT defines three key determinants of intention to use and additionally one determinant of actual behavior. The determinants for intention to use are performance expectancy, effort expectancy and social influence. The determinant for actual behavior is a set of facilitating conditions (Venkatesh et al., 2003). The determinants are moderated by gender, age, experience and voluntariness of use. The model explains 70% of the variance. Meta-analysis showed that in the longer run the performance expectancy was the strongest determinant of intention to use, whereas effort expectancy and social influence are weaker but also hold (Khechine et al., 2016). Further research has built on the UTAUT model. Umrani-Khan and Iyer (2009) added flexibility, interactivity and self-efficacy to their UTAUT based ELAM model. Balakrishnan (2016) has used that ELAM model together with the UTAUT model to apply its use to social media, the SMAM.

The SMAM model uses Performance, Effort/influence, Self and Communication Functionality as significant determinants to ‘predict 71% of intention to use social media for learning’ (Balakrishnan, 2016). This SMAM can be complemented by other ways of explaining the intention to use. Because this is such a large proportion it is believed that SMAM can be applied in a more general context too. This supports hypothesis 1, 4 and 6.

2.2 Privacy

In the rise of internet and big data privacy becomes more important and more vulnerable, so efforts need to be made for its preservation (Cohen, 2012). Solove (2002) argues that privacy should be conceptualized to be applied to real situations. Abstractions of privacy might not capture the concept fully. Therefore, in this research a privacy definition is used that was derived in the context of social media usage. The research into privacy by Trepte (2021) defines privacy as the individual's assessment of (a) the level of access to the individual in interactions, (b) the influence the individual has on the access from (a), (c) self-disclosure and (d) regulation of privacy. In a social media context point (b) is especially adjustable, because the goal of social media websites is to facilitate interactions between people.

Consumers are concerned about the collection and use of their personal data (Gutierrez, 2019). For instance, when privacy concerns are made salient, the targeting is ineffective (Goldfarb & Tucker, 2011). And consumers are becoming increasingly aware of the risks of data collection and usage (Bhargava & Velasquez, 2021; Muhammad et al., 2022). Research also shows the consequences of data collection by large companies can be detrimental (Janek, 2022).

In many studies a discrepancy between privacy awareness and preferences and privacy behavior is found. This is commonly referred to as the privacy paradox. In addition to personal attitudes and knowledge of risks, privacy decision making is influenced by many other factors. These factors cause decisions to be less private than their awareness or preferences would indicate (Acquisti & Grossklags, 2005; Barnes, 2006; Norberg, et al., 2007; Gerber et al., 2018; Willems et al. 2023; Dinev, 2024). Data privacy is well researched in the field of marketing (Martin & Murphy, 2017). The privacy paradox is also found in the social media context (Hayes et al., 2021; Bright et al., 2022; Ameen et al., 2022). The privacy calculus model states that consumers make a risk benefit analysis when deciding to make a purchase or use a service (Laufer & Wolfe, 1977; Culnan & Armstrong 1999). Gutierrez et al. (2019) offers this as an explanation for the existence of the privacy paradox. This supports hypothesis 2 and 3.

2.3 Intention to use social media determined by privacy

This paragraph is in support of hypothesis 7. Balakrishnan (2016) finds that the intention to use social media for learning is partially determined by the Performance, Effort/influence, Self and Communication Functionality. Privacy is not considered as a viable option. In a non-learning context where the benefits are more socially this might be different. Mariani et al., 2021 used privacy as a moderator in a data storage context while using the TAM, the effect of privacy was not significant in that context, but it is recommended that more research into privacy takes place. Privacy is described as a prerequisite for intention to use by Gutierrez et al. (2023). The effect of privacy is considered, but not as a determinant.

Privacy awareness and preferences have been used as a determinant in the context of purchasing products or to use a service. ‘Beliefs about privacy negatively affected attitude, but not as much as fear positively affected it.’ (George et al., 2021). The fear of missing out is different than the fear of a home break in, as George et al. (2021) researched. Privacy awareness and preferences might prove a determinant for intention to use social media. Similarly, Choi et al. (2023) used privacy as a determinant for continued use of travel apps. But the study could not definitively prove the triggering relation between perceived value of privacy protection and intention to use an app.

2.4 Confrontation with data collection

Direct confrontation with the collection and use of personal data might cause people to act more in line with their preferences. Consumers care about their privacy (Gutierrez, 2019; Bhargava & Velasquez, 2021; Muhammad et al., 2022). The direct confrontation might trigger consumers that are indeed conscious about their privacy. People will react according to their personal attitudes and knowledge of privacy risks (Acquisti & Grossklags, 2005). However, when privacy concerns are made salient consumer behavior can change (Goldfarb & Tucker, 2011). If a goal is made salient by a reminder and the risks outweigh the benefits different behavior might take place than without the reminder (Gravert, 2021). Similarly, when consumers are reminded of the privacy risks while making a social media account, this might lead to a different reaction than without the reminder. This supports hypothesis 5.

Chapter 3: Research methodology

This chapter discusses the data that is used to conduct the research and the two methods that were used in this research. Paragraph 3.1 describes what data that are collected by Instagram this serves as the foundation for the privacy policy summary in the experiment. Paragraph 3.2 describes the participants and the questions that were used to answer sub question 1, 2, and 3. In paragraph 3.3, the interviewing process is described as is the analyzing process. Paragraph 3.4 describes the data that was gathered in the survey and used to answer sub questions 4, 5, 6, and 7. Finally, the analysis of the survey data is discussed in paragraph 3.5.

3.1 Data collection

This paragraph is used to make the privacy policy summary for the experiment in the survey (Paragraph 3.3). In the recent years the data that is voluntarily placed on social media by consumers is strongly increasing and the use and collection are common practices for those social media companies (McCourt, 2018). Instagram states in their terms and conditions that they collect data and use personal data of their customers to effectively sell advertisement space on their service (Instagram n.d.). Instagram decides whom of their consumers are a good fit for advertisement wishes of third companies. They also track the effectivity of the advertisement. In their privacy policy Instagram specifies what data they exactly collect and use (Instagram, 2023, November 3rd).

The following data is collected by Instagram. Data provided by the consumer, audio provided by the consumer, camera data, face data, voice data, meta data of pictures and chats, app use, purchases and transactions, interactions with the consumer on Instagram, identifiers, meta data of the consumers photo storage, IP address, information on network and connectivity, performance of their service, cookies and similar technologies, data from partners, data from third parties, public sources. Instagram uses this data from all the devices the service is used on (Instagram, 2023, November 3rd).

Instagram also collects data about the device their service is used on. They collect: type of device, operating system, hardware and software details, brand and model, battery level, available storage, browser type, installed apps and file names and types, plug-ins. Additionally they collect what connections a device is making such as: GPS, Bluetooth-signals, nearby Wi-Fi access points, beacons and cell towers (Instagram, 2023, November 3rd).

In general, Instagram uses this data to: optimize advertisement effectivity, personalized advertisement, relevant topics, identify your network, define personality type. More specifically they use the data from device connections to localize the consumer for advertisement relevancy and confirm identity. They even do this when the location setting is disabled. In that case Instagram collect information on your location

using your IP-address, use of the service by the consumer and reported place of residence (Instagram, 2023, November 3rd).

Instagram uses cookies to track previous internet activity 'whether or not you are registered or logged in.' Amongst others Instagram uses these cookies to obtain a more complete profile of their users and to keep track of the performance of their advertisements (Instagram, 2023, December 12th).

In a comparing web post on invasive apps (Dimitrov, 2021), Instagram was found the most invasive. It was found to be tracking 62% of user's personal data. In a case study Janek (2022) explains why and how data collection by social media companies is compromising personal privacy. It is mentioned that especially the tracking of location has negative effects because it helps determine when a consumer is vulnerable for advertisements. Additionally, it is mentioned that most consumers don't read lengthy privacy and cookie policies, rendering them uninformed about these possible dangers.

Instagram collects a lot of data, almost everything that they can collect. They even collect data if the user tries to disable the data collection. In those cases, Instagram is able to collect such data or a proxy of this data (Instagram, 2023, November 3rd). They use the data to maximize advertisement sales. This can be achieved through finding the optimum between advertisement engagement and length of stay on the Instagram service.

3.2 Qualitative data

People with varying knowledge of privacy were interviewed, they all came from the Netherlands. Five interviews were conducted, they were held on Zoom or Microsoft Teams. In total nine invites for the interview were send out on LinkedIn and WhatsApp. The interviewer and participants were in their home office or workspace during the interviews. People from the age cohort that the main question focusses on (20 – 25) were interviewed, as well as people in their mid-fifties. 60% of the participants was male and 40% was female. The interviews took place on or between the 20th and 24th of May 2024. The interviews and the audio recording of the interviews were consented to by the participants by filling out the consent form (Appendix A). Some of the participants were friends or family from the Interviewer. The Interviews had a duration of 20 to 40 minutes. An overview of the participants is found in table 1.

Table 1

Characteristics of the participants in the interview

Participant	Age (years)	Gender	Main Occupancy	Education
1	22	Male	Research Analyst	BSc
2	54	Female	Quilt Artist	MSc
3	55	Male	Data Protection Officer	MSc
4	23	Male	Sailboat Rigger	1st year BSc
5	55	Female	Privacy Lawyer	MSc

To conduct the interviews an interview guide was developed and used (Appendix B). First the personal characteristics, hobbies and interests of the participant were asked. Then question relating to their social media use were asked. Thereafter the topic of privacy was mentioned, and the participants were asked about their idea on the value of privacy and on their behavior on privacy. A factsheet (Appendix C) on data collection was shown and subsequently the participants were asked if the information changed how they felt about the previous answers in the interview. Lastly, the interview was summarized by the interviewer and the participant were given the opportunity to comment on, add to, or subtract from their answers.

3.3 Qualitative research methodology

Interviews were used to gather the qualitative data that is necessary to answer sub questions 2 and 3. The interview started with a short introduction about the research and the interviewer and the on the first topic that would be discussed. Then some factual questions were asked and then the participant was eased into a more talkative position by asking them about their hobbies and interests. Then questions were asked about their social media use, their opinion on social media, as well as reasons for making their account. Then their opinion and behavior in relation to privacy was asked. Lastly, the participants were confronted with the data that is collected by a large social media company and they were asked if this changed the way they felt about their previous answers. In general, the participants were very cooperative and did not refuse to answer questions or to elaborate on their answers. There may be a selection bias regarding the people that consented to the interview relative to the ones that declined or ignored the invitation. The respondents were also frequently asked to expand on their answers and provide examples. In addition, the interviewer summarized the participants' responses which also led to specifications of those responses capturing the essence of what they meant more precisely.

The setting of the interview was informal and relaxed. The goal of the interview was to get insights into personal opinions of the participants. By creating an informal setting the participants were comfortable with sharing their opinions. All the interviews were in a videochat to ensure they were in their personal environment while keeping the possibility of looking at each other's facial expressions. This made it possible to dive deeper into the answers they gave.

All the interviews were audio recorded. The transcribe tool in Microsoft Word was used to automatically generate transcripts. Those were checked and revised by listening to the recording and confirming or improving the automatically generated transcripts. The transcripts were uploaded to Delve. Delve is a tool for thematic analysis, which allows researchers to assign parts of the transcripts to a certain code and group these codes into themes.

3.4 Quantitative data

The quantitative research contains an experiment. It was carried out by sending a survey to people from the age of 20 to 25 in the Netherlands. The experiment contained one treatment group and one control group. The treatment group was shown a privacy policy summary, where the control group saw the regular Instagram-like interface to make a social media account (Appendix E). The name and the logo of the fictional social medium were generated with ChatGPT 3.5 and the AI logo creator of Design.com. The description of the fictional social medium was made by the researcher based on Instagram and the results of the qualitative research. The survey was primarily sent to people withing the personal network of the researcher and was send onwards to other people within their networks. 104 people filled out the survey, of them 87 people that met the condition of living in the Netherlands and being between the age of 20 and 25. The survey was send out in the week of Monday the 27th of May 2024. The sample was collected on Monday 3rd of June 2024. The respondents were on average 22.9 years old and were 49.4% male and 49.4% female, with one person not identifying with either of the genders. A strong majority of respondents are from Zuid-Holland and have a university education. The majority of respondents have multiple social media accounts, most common are Instagram, WhatsApp, Snapchat and LinkedIn. Respondents mostly indicate that the decision to use of social media is their own, so they use the services voluntarily and do not feel forced (Appendix F).

The 5 variables *Performance*, (P) *Self*, (S) *Communication Functionality*, (CF) *Influence/Effort* (IE) and *Privacy Awareness and Preferences* (PAP) were used to determine the intention to make a social media account. They were found in the study of Balakrishnan (2016) and Choi et al. (2023). All 5 variables were asked with 5 questions that needed to be answered on a 7-point Likert scale (Appendix E, average answers Appendix F). The questions were adjusted and complemented with findings from the qualitative part of this research (Paragraph 4.2). Additionally, the results were checked for possible moderation by Gender, Age, Experience, Voluntariness of Use (Venkatesh, et al., 2003) Level of Education to be able to diversify the group of respondents. Age and Place of Residence were asked to determine whether the respondent was part of the studied subgroup (Appendix E). The experiment was operationalized by showing 50% of respondents a privacy policy summary and 50% the regular interface.

Table 2
Summary statistics of the experiment

	Mean	Median	Std. Dev.	N
Treatment	3,20	3	1,593	44
Control	4,51	5	1,653	43

Note: [a] N = 87

3.5 Quantitative research methodology

The Qualtrics program that was used randomly divided the respondents into two groups. The treatment group who will see the privacy policy summary and the control group who will not see the summary. Similar to the approach used by Paul & Rana (2012) the results from the survey will be analyzed using multivariate analysis: factor analysis and linear regression analysis.

First an independent sample t-test will be used to check if the treatment group significantly differs from the control group in their answer on the experiment. Then an independent sample t-test will be used to check if the two test groups were statistically not different.

Subsequently, a factor analysis will take place to determine what questions carry the same loading. The principal component analysis is then used to extract the factors and the orthogonal rotation method Varimax is used to make the result more interpretable. The Eigen Value needs to be at least 1.0. The factor loadings were cut off below 0,7 to maintain coherent factors. The KMO and Bartletts-test are used to check the sample and its validity. The Cronbach's Alpha is used to check if the factors indeed carry the same meaning (hypothesis 4).

An independent sample t-test is also used to see if the respondents who were in the treatment group answered differently on the privacy awareness and preference questions compared to the control group (hypothesis 5). A linear regression is performed to predict the Intention to Make a Social Media Account (IMSMA) with the factors from the factor analysis controlled for the two experiment groups. The R^2 are used to determine what combination is the most robust (hypotheses 6 & 7).

$$IMSMA_i = \beta_0 + \beta_1 \boldsymbol{Factors}_i + \beta_2 \boldsymbol{Dummy\ Experiment}_i + \varepsilon_i$$

The assumptions for linear regression: linearity, independence, homoscedasticity, normality of residuals, and no perfect multicollinearity are checked. The results seem to be pretty linear. The Durbin Watson test was performed to check independence. A scatterplot of the unstandardized residual is run to confirm homoskedasticity. The normality of the residuals is checked with a Shapiro-Wilk test. Multicollinearity is checked with a correlation matrix.

Chapter 4: Results & Discussion

This chapter presents the results of the research and discusses them. First the qualitative results are presented and discussed and secondly the quantitative results. This is followed by the empirical framework that summarizes the most important findings.

4.1 Results of the qualitative research.

From the interviews a number of themes we discovered (Table 3). In this part these themes will be categorized and presented.

Sub question 1: What do Dutch residents from various ages consider when making a social media account, such as Instagram?

While making and using a social media account many factors play a role. The positive side of using a social media account is mostly determined by the information that can be gathered, how well it helps organizing one's social life, what leisure it can provide and, in some instances. Also, the use for professional purposes can contribute to the value of social media. It is also found that people often start using a social medium, if their peers are using it as well. 'I think it's also a situation of everyone is in WhatsApp. So, you start using it'.

The opposite is also found in the interviews: 'For instance, if a lot of people would go away from Instagram because they were concerned, then I might also leave because then it's not as interesting anymore'. Other concerns while making an account and using social media are the social unsafety that comes from fake news and negative comments. Another prevalent concern is getting caught in the app, in the interviews it was mentioned that people spend more time using the services than they intended. Other downsides are perceived pressure and a negative self-image form only seeing the positives in other users' lives.

Sub question 2: What are possible reactions to confrontation with the data collection practices from social media companies?

Data collection stopped people from using Facebook and also served as a reason for one of the participants not to join Facebook even when it means missing out on some social groups. 'So that kept me from ever going on Facebook, and sometimes I do miss out on things, on kind, on certain discussion groups that I know exist on Facebook.'

When confronted with the data collections from the Data Collection Factsheet three different reactions were distinguished. Firstly, some participants were unimpressed. It was indicated that they were more or less aware about the data collection practices, but that they did not change their behavior based because of this knowledge. Secondly, some participants were caught off guard by the information. They were surprised

about the information and felt fooled and unhappy, because of the data collection practices. One participant indicated this type of information was the reason they stopped using that social medium and another participant indicated this type of data collection practices were part of the reason they never got on the social medium. Thirdly, some participants were passively discontent. They were surprised by at least some of the information. But the benefits of using social media would not out weight the perceived risk of the data collection. They also indicated they were unaware of ways to prevent this data collection or felt unable to change the situation.

Sub question 3: How do privacy awareness and preferences influence the intention to make a social media account?

In most interviews the topic of privacy was discussed after it had been introduced by the interviewer. They rarely mentioned it as a reason to discontinue their social media use or as a reason to not start using social media. Privacy was not dominant in the answers to the questions relating to reasons to make a social media account. And even after being confronted with the Data Collection Factsheet some participants indicated they would still use social media.

All the participants indicated that they control what level of information they shared on their social media: some would draw the line at their place of residence, some at information that could be used for identity fraud and some preferred not to use social media when talking about sensitive topics with close contacts. The tendency of the answers leaned to using social media even though it has severe privacy issues. A key argument to make this acceptable for themselves, is through controlling what information they wanted to share, and some participants indicated they would opt for more privacy of it was easier. Nonetheless, the balance between using social media and privacy is in favor of using social media. The benefits outweighed the privacy concerns.

Table 3
Theme overview per question category

	Confrontation	**Privacy in Social Media**	**Social Media**	**Privacy**
Themes	unimpressed	Against data collection	information gathering	Relativity
	caught off guard	counter code to ↑	Social unsafety	Opt-in privacy
	Passively discontent.		downside to social media	Control
			leisure	Privacy activist
			caught in the app	
			professionality	
			organizing social life	

4.2 Discussion of the qualitative research.

Hypothesis 2 cannot be rejected. Performance, Effort/influence, Self and Communication Functionality were all elements that come forward in the interviews. However, the focus was slightly different. The main reason to use social media and to make a social media account was to interact with large user base of peers (p. 20).

Hypothesis 3 cannot be rejected. There was a calculus effect in people discarding social media services because of data collection. The outcome differed, some participants stopped using or never started using because of the data collection, some participants did not care about the data collection and some participants disliked the data collection but would not change their behavior based on that. Those were the tree possible reactions to the data collection practices (p. 20-21).

Hypothesis 4 should be rejected. The participants were not really concerned about their privacy while making a social media account. Only after a reminder they recalled some experiences where they consider privacy in their decision to use social media. It cannot be said that it played an important role I their initial decision to make a social media account. The benefits outweighed the concerns for privacy, confirming the theory of Laufer & Wolfe (1977) and Culnan & Armstrong (1999). The results also give a contra indication to the fact that people are substantially concerned with their privacy (Gutierrez, 2019; Bhargava & Velasquez, 2021; Muhammad et al., 2022) (p. 21).

4.3 Results of the quantitative research.

The mean score of the answer of the treatment group to the question whether they would make a social media account was 3.20. The mean score of the control group on the same question was 4.51 and the difference was 1.307. Using Levene's test equal variance could be assumed. The difference was significant on the <0,001 significance-level. So, the respondents had a strong negative reaction to the reminder on data collection. In the results of the quantitative research, it will be explained that this is a reaction to the reminder and that the reminder does not affect the privacy awareness and preferences of respondents.

Sub question 4: Are Performance, Effort/influence, Self and Communication Functionality also applicable in a general social media context as explaining factors?

The Kaiser-Meyer-Olkin Measure of Sampling adequacy (KMO) was 0,572, The Bartlett's Test of sphericity was significant ($\chi^2(300) = 847$; $p<0.001$). Bartlett's test suggests the data is suitable for factor analysis, the KMO measure suggest it is miserable, however not unacceptable. Nine factors were determined based on the factor analysis (Appendix G, Table 10). Each factor was named based on the underlying questions that carried the same meaning. The factors were named: Necessary Communication (NC), Ease of Use (EU), Two Way Communication (TWC), Feeling Enabled (FE), Popularity (Pop), Privacy Awareness (PA), Privacy Preferences (PP), Self-Image (SI) and One Way Communication (OWC). Except for NC all the combined factors had appropriate Cronbach's Alphas (Table 4).

Table 4
Summary statistics and Cronbach's Alpha of factors

	Based on questions	Mean	Median	Std. Deviation	Cronbach's Alpha
NC	P4, P1	4,9425	5,5	1,3039	0,47
EU	IE5, IE4, IE3	5,1456	5,3333	1,11335	0,86
TWC	CF3, CF1, CF2	4,2835	4,3333	1,08135	0,70
FE	S2, S3, S1	5,8352	6	0,8397	0,77
Pop	IE2, IE1	5,1609	5,5	1,39662	0,82
PA	PAP1, PAP2	2,1322	2	1,22108	0,87
PP	PAP4, PAP5	5,2644	5,5	1,04503	0,73
SI	S4	6,023	6	0,8209	NA
OWC	CF5	5,7471	6	0,93035	NA

Sub question 5: Does a reminder on privacy increase the privacy awareness and change preferences?

The differences between the control and experiment group were very small. The largest difference in average answers between the two groups was 0,27 on the 7 point Likert-scale. They also were not significant (Appendix G, Table 11). So, no difference in the answers on the privacy awareness and preferences between the control and experiment group can be proven. Therefore, it cannot be said that the reminder on data collection has effect on the privacy awareness and preferences.

Sub question 6: Do Performance, Effort/influence, Self and Communication Functionality explain the intention to make a social media account?

Sub question 7: Do privacy awareness and preferences explain the intention to make a social media account?

Table 5
Linear regression models

model	Constant	NC	EU	TWC	FE	Pop	PA	PP	SI	OWC	Dummy Experiment	adjusted R^2
1	1,821	0,042	0,014	-0,187	0,257	0,361	0,013	-0,184	0,170	-0,031	-1,420	0,170
2	2,240	0,054	-0,084	-0,258	0,329	0,292**	-0,029	-0,105	0,097	-0,090		0,110
3	1,774			-0,279	0,336	0,255**						0,064
4	3,674***			-0,286*		0,272**						0,049
5	1,907			-0,180	0,261	0,368**		-0,201	0,177		-1,421***	0,210
6	3,021*			-0,183	0,278	0,37**		-0,231			-1,398***	0,213
7	2,312				0,282	0,358**		-0,232			-1,459***	0,209
8	2,904***					0,325**					-1,440***	0,190

Note: * p < ,1, ** p < ,05 *** p<,001.

Several combinations of factors and controls have been tried. The model with the highest adjusted R^2 is model 6. In model 6, 21,3% of IMSMA is explained by TWC, FE, Pop and PP, controlled for the experiment. However, only Pop and the control for the experiment are statistically significant. The Constant is significant at a more lenient level. The overall best model is model 8, it still explains 19% of IMSMA and is statistically significant. Model 8 looks like this:

$$IMSMA_i = 2{,}904 + 0{,}325 * Popularity_i + -1{,}440 * Dummy\ Experiment_i + \varepsilon_i$$

If people indicate that their social media use is driven by its popularity the answer to whether they would make a social media account increases by 0,325 on average. If people were in the experiment group, they would the answer to whether they would make a social media account is 1,440 lower on average. 4,00 means they did neither agree nor disagree to the statement that they would make an account. Anything lower is a negative response, anything higher is a positive response. The assumptions of linear regression mostly hold. However, it cannot be said that the residuals are following a normal distribution. And multicollinearity might exist between Popularity and Privacy Preferences in model 6 (Appendix H)

4.4 Discussion of the quantitative research.

Hypothesis 4 should be rejected, even though the factors found in this study corresponded with the factors in the study of Balakrishnan (2016). Self is partially covered by Feeling Enabled and Self Image, Performance is partially covered by Necessary Communication, Influence/Effort are partially covered by Ease of Use and Popularity and Communication Functionality are partially covered by One Way Communication and Two Way Communication. The determinants of Balakrishnan (2016) were not directly retrieved from the survey data. The KMO scores were too low. Leaving out questions on Performance and Privacy Awareness and Preferences (PAP) did not yield much higher KMO scores, nor did the resulting factors yield more significant results than is presented in paragraph 4.3. This control was made because the Survey Questions on Performance were adapted to the general context more rigorously and PAP was added in this research. This adaptation of the questions from Balakrishnan (2016) might have caused the determinants to be less clearly separated. Another explanation can be that the determinants might need to be refined more to the general context of the intention to make social media than was done in this research by using the qualitative research and other literature. Additionally, the amount of respondents in this study was much lower than the amount of respondents in the study of Balakrishnan (2016). More respondents might lead to clearer correlation patterns and lower the sampling error. (p. 23).

Hypothesis 5 should be rejected. There were no statistical differences in answers on questions about Privacy Awareness and Preferences between the group that got the privacy policy summary and the control group. They did answer differently on the question whether they would make an account. This indicated that the privacy policy summary served as a reminder that changed the direct behavior but did not change the short term convictions of the respondents. The effect of the reminder could have run out by the time the respondent reached the questions on privacy. Alternatively, the reminder only had a direct effect on the immediate decision and did not make the goal of well protected privacy more salient. It is claimed the latter does happen (Gravert, 2021) making the former explanation more likely. The measuring of the Privacy Awareness and Preferences could have been faulty or taken place too late (p. 24).

Hypothesis 6 and 7 can both be rejected. Performance, Effort/influence, Self and Communication Functionality and Privacy Awareness and Preferences did not prove to be predictors of the Intention to Make a Social Media Account (IMSMA). In model 6, Two Way Communication and Privacy Preferences had a negative effect and Feeling Enabled and Popularity had a positive effect the control for being in the experimental group proved to be the largest effect and was negative. It was unexpected that Two Way Communication had a negative effect, because it would be a benefit of social media use in the study of Balakrishnan (2016). Model 8 was the only fully significant model. It Predicted that Popularity was the determinant of IMSMA and that the experiment heavily influenced IMSMA.

A possible explanation for the difference in results between the study of Balakrishnan (2016) and this study might be the set-up of the experiment where the social media account respondents could indicate to make, was fictional. Because the fictional component the description of the social media where people could make an account for might have influenced their decision. In this description there was a focus on popularity of the social media. Therefore, the results might be stronger for Popularity. The normality of the residuals also indicates that there might be a negative factor missing to determine the Intention to Make a Social Media Account. This negative factor could be that the social medium is fictional.

A reason that some factors were not significant might have been the relatively small group of respondents for the large number of questions and answering options there were. This is likely for the factors that leaned to the significant side. For the factors that did not lean to the significant side an explanation might be that the adaptation of the questions did not align enough with the factors (p. 24).

4.5 Empirical framework

The qualitative research gave insight into the reasons for social media use and privacy concerns that largely corresponded with the theoretical framework. It only contradicted recent studies where privacy concerns were stronger than the intention to use. It is likely that this was caused by the different context those results were in.

The quantitative research mostly did not correspond with the expected results. The expectation that there were multiple groups of social media users, was true. But the determining factors (Balakrishnan, 2016) and the causal relation between those factors and the Intention to Make a Social Media Account were rejected in this study. This might be because the experiment was too fictitious, and the description of the social media might have leaned in favor of certain factors. This is supported by a skewedness to the left of the residuals. The adaptation of the questions to form the factors also might have caused a discrepancy between the factors of Balakrishnan (2016) and the factors found in this study. Lastly the sample might have been too small to yield significant results.

It cannot be said that the reminder increases the privacy awareness or preferences, but the confrontation with the data collection clearly influences the decision to make a social media account.

Chapter 5: Conclusion, Implications, Limitations & Recommendations

5.1 Conclusion

In this study the importance of privacy awareness and preferences is studied in the context of social media use. Previous research has widely researched social media use in various contexts. This study focused on the moment before using social media, when the decision to make an account is made. Privacy awareness and preferences have been used as determinant but never proven successful. However, privacy is found as a concern for people when using social media. Therefore, this research tries to find when this concern might result into action. Because it is found that reminders can alter behavior, the research question is:

How will Dutch residents from 20 – 25 years react to confrontation with data collection practices, when making a fictional social media account?

The insights in the exact concerns and benefits of social media use and the views on privacy were found by interviewing privacy aware and unaware individuals. Then a survey was sent out to Dutch residents between the age of 20 and 25. The treatment group got a reminder about the data collection and a sign-up page from a fictional social medium when they were asked if they would make an account, the control group only saw the sign-up page. The interviews showed that people did care about their privacy their privacy when confronted with the data collection but did not care too much about privacy when using social media. The survey confirmed that people would strongly react to a reminder about data collection practices, but that it did not change their privacy awareness and preferences.

This study concludes that Dutch residents from 20 – 25 years have a strong negative reaction to confrontation with data collection practices when making a fictional social media account. This is an immediate effect of the reminder and changes nothing in their privacy awareness and preferences. The decision to make a social media account is not caused by privacy awareness and preferences. The privacy awareness and preferences cause people to have a strong negative reaction to the confrontation with data collection.

5.2 Implications

This implies that an effective way to prevent data collection by social media companies would be to make such a reminder mandatory when signing up for a social media account. Then people would either not make an account, or the data collection has to be limited by the social media companies. This also implies that it is in the best interest of social media companies to give as little information as possible about their data collection because their user base would grow less or even decline.

5.3 Limitations

Theoretical Framework:

The theory of Solove (2021) on the privacy paradox has not been used in this study. Solove claims that the privacy paradox does not exist on the basis that privacy has many dimensions. Therefore, a preference in one context of privacy cannot contradict a behavior in another context of privacy, even though it seemingly does look like preference and behavior contradict each other. This is a recent theory that does not form the consensus in academic field. Also, this contradiction is largely solved in the privacy calculus theory, where privacy preferences are still one dimensional. Nonetheless, it is an interesting theory that might help to uncover complex privacy behavior and preferences (as in Chapter 2.2).

The determinants, which were borrowed from Balakrishnan (2016), did not prove to be applicable in the general context of social media. This is despite of the additional literature and the qualitative research that was used to adjust the determinants. More empirical research, be it qualitative or quantitative, can help to improve the determinants and make them better applicable to the context of making a social media account.

Research Methodology:

For the interviews, people with varying knowledge about privacy were interviewed. This resulted in interesting insights, but the knowledge of privacy was quite extensive. More research into what the 'common person' thinks about privacy, might help further refine the determinant of privacy awareness and preferences.

Ideally the experiment would have been carried out in a non-fictional situation. In that situation, the consequences of not being on a certain social medium would have been real. Also, the experiment would gain a time dimension. Participants would continuously be able to make an account for the social medium, whereas in this experiment the participants only had one moment to decide whether they would make an account or not.

Results & Discussion:

The study shows a strong negative reaction to confrontation with data collection practices. Although the privacy policy summary was constructed with real data collection practices, the effect was very strong. A smaller reminder using only one aspect of the data collection, or less severe practices might be interesting. Different thresholds for people to react could be identified when a wide array of privacy reminders is used.

5.4 Recommendations

For future research it is recommended that the limitations are taken into account. It will be especially interesting to continue research in a real life setting to exclude side effects from a fictional situation. But it would already be interesting to do the experiment in a less fictional setting. This is more realistic, because it might prove difficult to find a large social media company willing to cooperate.

Future research could also focus on the determinants that have influence on the intention to make a social media account. The determinants of Balakrishnan (2016), did not satisfactorily explain the intention to make a social media account. As research advances, there will be more insights that would have been useful for this study. More quantitative studies and intensive qualitative research might help to further refine the possible determinants for the intention to make a social media account.

A last angle for future research could be the privacy reminder. A wider range of privacy reminders in terms of amount and type of information will help disclose the threshold when people will react to such a reminder. This would further explain what people value in privacy. The reaction could then also be diversified from a binary choice between making or not making an account, to multi-option choice where they can adjust what data is collected. This will help explain privacy preferences even further.

References

Academic papers

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE security & privacy*, 3(1), 26-33. http://doi.org/10.1109/MSP.2005.22

Ameen, N., Hosany, S., & Paul, J. (2022). The personalisation-privacy paradox: Consumer interaction with smart technologies and shopping mall loyalty. *Computers in Human Behavior, 126*, 106976, 1-12. https://doi.org/10.1016/j.chb.2021.106976

Arulogun, O. T., Akande, O. N., Akindele, A. T., & Badmus, T. A. (2020). Survey dataset on open and distance learning students' intention to use social media and emerging technologies for online facilitation. *Data in brief*, 31, 105929, 1-8. https://doi.org/10.1016/j.dib.2020.105929

Balakrishnan, V. (2016). Key determinants for intention to use social media for learning in higher education institutions. *Universal Access in The Information Society*, 16(2), 289–301. https://doi.org/10.1007/s10209-016-0457-0

Barnes, S. (2006). A privacy paradox: Social networking in the United States. *First Monday*, 11(9). http://dx.doi.org/10.5210/fm.v11i9.1394

Barth, S., & De Jong, M. D. (2017). The privacy paradox–Investigating discrepancies between expressed privacy concerns and actual online behavior–A systematic literature review. T*elematics and informatics,* 34(7), 1038-1058. https://doi.org/10.1016/j.tele.2017.04.013

Bhargava, V. R., & Velasquez, M. (2021). Ethics of the attention economy: The problem of social media addiction. *Business Ethics Quarterly,* 31(3), 321-359. https://doi.org/10.1017/beq.2020.32

Bright, L. F., Logan, K., & Lim, H. S. (2022). Social media fatigue and privacy: An exploration of antecedents to consumers' concerns regarding the security of their personal information on social media platforms. *Journal of Interactive Advertising*, 22(2), 125-140. https://doi.org/10.1080/15252019.2022.2051097.

Cohen, J. E. (2012). What privacy is for. *Harvard Law Review*, 126, 1904-1933. https://heinonline.org/HOL/P?h=hein.journals/hlr126&i=1934

Choi, K., Wang, Y., Sparks, B. A., & Choi, S. M. (2023). Privacy or security: does it matter for continued use intention of travel applications?. *Cornell Hospitality Quarterly*, 64(2), 267-282. https://doi.org/10.1177/19389655211066834

Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization science*, 10(1), 104-115. https://doi.org/10.1287/orsc.10.1.104

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340. https://psycnet.apa.org/doi/10.2307/249008

Dinev, T. (2014). Why would we care about privacy?. *European Journal of Information Systems*, 23, 97-102. https://doi.org/10.1057/ejis.2014.1

George, J. F., Chen, R., & Yuan, L. (2021). Intent to purchase IoT home security devices: Fear vs privacy. *PLoS one,* 16(9), e0257601, 1-14. https://doi.org/10.1371/journal.pone.0257601

Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, 77, 226-261. https://doi.org/10.1016/j.cose.2018.04.002

Goldfarb, A., & Tucker, C. (2011). Online display advertising: targeting and obtrusiveness. *Marketing Science*, 30, 389–404. https://doi.org/10.1287/mksc.1100.0583

Gravert, C. A. (2021). Reminders as a tool for behavior change. *SSRN Electronic Journal*. https://dx.doi.org/10.2139/ssrn.3888238
(A version of this article is in Mazar N and Soman D., *Behavioral Science in the Wild, Toronto, Canada: University of Toronto Press*.)

Gutierrez, A., O'Leary, S., Rana, N. P., Dwivedi, Y. K., & Calle, T. (2019). Using privacy calculus theory to explore entrepreneurial directions in mobile location-based advertising: Identifying intrusiveness as the critical risk factor. *Computers in Human Behavior*, 95, 295-306. https://doi.org/10.1016/j.chb.2018.09.015

Gutierrez, A., Punjaisri, K., Desai, B., Alwi, S. F. S., O'Leary, S., Chaiyasoonthorn, W., & Chaveesuk, S. (2023). Retailers, don't ignore me on social media! The importance of consumer-brand interactions in raising purchase intention-Privacy the Achilles heel. *Journal of Retailing and Consumer Services*, 72, 103272, 1-39. https://doi.org/10.1016/j.jretconser.2023.103272

Hayes, J. L., Brinson, N. H., Bott, G. J., & Moeller, C. M. (2021). The influence of consumer–brand relationship on the personalized advertising privacy calculus in social media. *Journal of Interactive Marketing*, 55(1), 16-30. http://dx.doi.org/10.1016/j.intmar.2021.01.001

Janek R (2022). The Dark Side of Social Media Data Collection and Retention. *Journal of Applied Management Accounting Research*. 20. 1-7.

Khechine, H., Lakhal, S., & Ndjambou, P. (2016). A meta-analysis of the UTAUT model: Eleven years later. *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration,* 33(2), 138-152. https://doi.org/10.1002/cjas.1381

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, 43(6), 740-755. https://doi.org/10.1016/j.im.2006.05.003

Laufer, R. S., & Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of social Issues*, 33(3), 22-42. https://doi.org/10.1111/j.1540-4560.1977.tb01880.x

Mariani, M. M., Styven, M. E., & Teulon, F. (2021). Explaining the intention to use digital personal data stores: An empirical study. *Technological Forecasting and Social Change*, 166, 120657, 1-51. https://doi.org/10.1016/j.techfore.2021.120657

Martin, K. D., & Murphy, P. E. (2016). The role of data privacy in marketing. *Journal Of The Academy Of Marketing Science*, 45(2), 135–155. https://doi.org/10.1007/s11747-016-0495-4

Muhammad, S. S., Dey, B. L., Alwi, S. F. S., Kamal, M. M., & Asaad, Y. (2022). Consumers' willingness to share digital footprints on social media: the role of affective trust. *Information Technology & People*, 36(2), 595-625. https://doi.org/10.1108/itp-10-2020-0694

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs*, 41(1), 100-126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

Paul, J., & Rana, J. (2012). Consumer behavior and purchase intention for organic food. *Journal of Consumer Marketing*, 29(6), 412–422. https://doi.org/10.1016/j.jretconser.2017.06.004

Schumann, J. H., Wangenheim, F. V., & Groene, N. (2014). Targeted online advertising reciprocity appeals to increase acceptance among users of free web services. *Journal of Marketing*, 78, 59–75. https://doi.org/10.1509/jm.11.0316

Solove, D. J. (2002). Conceptualizing privacy. *California Law Review*, 90, 1087-1155. https://doi.org/10.2307/3481326

Solove, D. J. (2021). The myth of the privacy paradox. *George Washington Law. Review*., 89, 1-51. https://dx.doi.org/10.2139/ssrn.3536265

Trepte, S. (2021). The social media privacy model: Privacy and communication in the light of social media affordances. *Communication Theory*, 31(4), 549-570. https://doi.org/10.1093/ct/qtz035

Umrani-Khan, F., & Iyer, S. (2009). ELAM: a Model for Acceptance and use of e-Learning by Teachers and Students. *In Proceedings of the International Conference on e-Learning, Institute of Technology Bombay, Mumbai, India*, 475-485.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478. http://dx.doi.org/10.2307/30036540

Willems, J., Schmid, M. J., Vanderelst, D., Vogel, D., & Ebinger, F. (2023). AI-driven public services and the privacy paradox: do citizens really care about their privacy?. *Public Management Review*, 25(11), 2116-2134. http://dx.doi.org/10.1080/14719037.2022.2063934

Yuan, D., Rahman, M. K., Gazi, M. A. I., Rahaman, M. A., Hossain, M. M., & Akter, S. (2021). Analyzing of User Attitudes Toward Intention to Use Social Media for Learning. *SAGE Open*, 11(4), 1-13. https://doi.org/10.1177/21582440211060784

Institutions, Websites & Newspapers.

AP (n.d.). Het belang van privacy.

Retrieved 25 June 2024, from:

https://www.autoriteitpersoonsgegevens.nl/themas/basis-avg/privacy-en-persoonsgegevens/het-belang-van-privacy

Dimitrov I. (2021) Invasive Apps.

Retrieved 25 June 2024, from:

https://www.pcloud.com/nl/invasive-apps

Instagram (n.d.) Terms and Conditions.

Retrieved 25 June 2024, from:

https://help.instagram.com/581066165581870?ref=dp&helpref=faq_content

Instagram (2023, November 3) Privacy policy.

Retrieved 25 June 2024 from:

https://privacycenter.instagram.com/policy

Instagram (2023, December 12) Cookie policy.

Retrieved 25 June 2024, from:

https://privacycenter.instagram.com/policies/cookies/

McCourt (2018, April 3) Social Media Mining: The Effects of Big Data In the Age of Social Media.

Retrieved 25 June 2024, from:

https://law.yale.edu/mfia/case-disclosed/social-media-mining-effects-big-data-age-social-media

OHCHR. (2022). Privacy and data protection: Increasingly precious asset in digital era says UN expert.

Retrieved 25 June 2024, from:

https://www.ohchr.org/en/press-releases/2022/10/privacy-and-data-protection-increasingly-precious-asset-digital-era-says-un

Appendices

Appendix A: Information and Consent Form

Information and Consent Form: Bachelor's Thesis on Social Media Use

Introduction

Dear (Potential) Participant,

My name is Emil, and I am an economics and business economics student at Erasmus School of Economics. I am conducting research for my thesis. My study is about social media use. This information sheet will provide the most important details of my study. If you have any questions after reading this sheet, please feel free to ask any questions. The consent form in the end of this document needs to be filled out before taking part in the research.

Research topic

My study aims to find out what important factors are for making a social media account. The answers in this survey will serve to give an academic foundation to some of the questions that will be asked in a survey format. Your participation is requested to find a diverse group of respondents.

The Interview

An interview of roughly 15 to 20 minutes. The interview can take place in a café, the Erasmus University Rotterdam, your home, or online through Zoom or Teams. With your permission, the audio of the interview will be recorded. You are free to choose to answer a question or not. If requested, parts of the interview can be redacted or amended.

Participation in this study is completely voluntary and can be withdrawn at any time. If you decide to withdraw your participation, no explanation is necessary.

What will be asked

I will ask for contact information, as well as general background information such as name, age, study/occupation, hobbies, and interests. Then questions will be asked about making a social media account.

Data protection and handling

- Your data will be stored securely.
- Only the principal researcher has access to the not anonymized data. Supervisors will have access to the anonymized data.
- All data will be securely stored.
- Transcriptions of the audio recording will be made, please note that the recording will start after the questions on your name and place of residence.
- If direct personal data such as name and place of residence will be recorded the will be redacted form the transcript.
- The data will be used to design survey questions for the second part of my thesis.
- Answers might be directly used in the study. If this contains information that might identify you will be asked for prior consent again.

Taking the previous into account, your identity might be inferred from your answers. Effort will be taken to minimize this risk.

Data storage period

The transcripts and audio recordings will be stored until 12 months after the completion of the thesis.

In case of questions after signing the consent form.

If you have any questions about the study or your privacy rights, such as accessing, changing, deleting, or updating your data, please contact me through the below email address.

Name: Emil de Jong

Email: 532158ej@student.eur.nl

Do you have a complaint or concerns about your privacy? Please email the Data Protection Officer (fg@eur.nl) or visit www.autoriteitpersoonsgegevens.nl. (T: 088 - 1805250)

Withdrawing from the study

If you want to withdraw from the study this is possible at any moment until the data has been anonymized. Anonymizing will take place in the week after the interview.

Declaration of Consent

I have read an understood the information sheet. If I had any questions I have asked them, and the answers were adequate.

By signing this form, I:

- ☑ Consent to participate in this research;
- ☑ Consent to the use of my personal data
- ☑ Confirm that I am at least 18 years old;
- ☑ Confirm that I understand that participating in this research is completely voluntary and that I can stop at any time;
- ☑ Confirm that I understand that my data will be anonymized for publication, educational purposes and further research;

Check the boxes below if you consent to this.

Data [about specify]

☐ I consent to the collection, use and retention of the following data: My thought process when making a social media account.

Audio recording

☐ I consent to [the interview] being audio recorded.

Name of participant:

Participant's signature: Date:

Appendix B: Interview Guide

1. Introduction:

This interview will take about 15 minutes, the topic is mainly social media and related preferences.

1.1 What is your highest level of education, and where did you follow that education?

1.2 What is your age?

1.3 Where do you live?

1.4 What kind of job do you have or what is your daily occupation?

1.5 What are your hobbies?

1.6 What are you interests, for instance when reading a newspaper or browsing the web what would you read about?

2. Social media general

2.1 What types of social media do you use?

2.2 How much do you use it?

3. Social media specific

3 Per type of social media:

3.1 Do you remember making your account?

3.2 What were your reasons to make a social media account?

3.3 What do you see as benefits of social media?

3.4 What are your concerns when using social media

Summary by interviewer

4. Privacy:

Now we will talk about a different topic, to not influence your previous answers I will tell you now what it is: Privacy.

4.1 What does privacy mean to you?

4.2 How much do you value privacy?

4.3 Do you take actions on those preferences?

Summary by interviewer

5. confrontation with data

Data Collection Factsheet is shown.

Does this information make you feel different about your previous answers? Why?

6. Additional questions

What question do you feel I haven't asked but you do want to elaborate on?

Are you happy with your answers or would you like to add anything?

Summary by interviewer

Interview Techniques (Sonja Wendel and Emilie Huber in the subject Qualitative Research in Marketing)

Probing questions as needed (apply to all questions):

What do you mean exactly?

Can you provide an example?

Can you describe what you mean?

Can you provide some more explanation?

Can you tell a bit more about this?

How do you deal with this?

Summarize what has been said to see if you captured it all and well after every section by asking for instance:

Did I hear you say…?

Did I understand you when you said…?

Did I hear you correctly when you said…?

Is this what you said…?

Keeping track of the interview:

Thank you for sharing this information me, but I would like to return to an earlier question / something you said…

Thank you, but we need to continue now with…

Appendix C: Data Collection Factsheet

https://www.nytimes.com/2018/04/11/technology/facebook-privacy-hearings.html

https://www.theguardian.com/commentisfree/2018/mar/28/all-the-data-facebook-google-has-on-you-privacy

https://allaboutcookies.org/what-data-does-facebook-collect

What data does Facebook collect outside of the website?

Facebook gathers significant amounts of data from outside its own services to enhance and personalize user experiences. Here's an overview of the external data Facebook collects:

Data from devices

Facebook collects a variety of device-related information, such as the following:

- **Device attributes:** Operating system, hardware versions, battery level, storage space, and browser details.
- **Device operations:** User behaviors like application usage and mouse movements.
- **Identifiers:** Unique device and application identifiers.
- **Device signals:** Nearby Bluetooth signals and information about Wi-Fi access points.
- **Settings data:** Information such as GPS location, camera, and photo library access enabled by the user's device settings.
- **Network information:** Data like mobile operator, internet service provider, IP address, and mobile phone number.

Information from partners

Facebook also receives data from third-party advertisers, app developers, and publishers. This includes:

- **User activity data:** Information on device usage, websites visited, purchases, and viewed ads gathered through tools like Facebook Pixel.
- **Interactions with services:** Data on how users interact with partners' services, regardless of whether they have a Facebook account or are logged in.

Cookies and similar technologies

Facebook uses cookies to track user activities both on and off its platforms to support advertising and other site functions, enhancing content personalization and measuring ad effectiveness.

This data helps Facebook not only to tailor ads more effectively but also to improve the overall security and functionality of its services.

Appendix D: Thematic analysis

Table 6

Thematic analysis with themes, codes and quotes

question category	theme name	codes	quotes
confrontation	unimpressed	unsurprised	There's nothing we haven't heard before on the news, right?' I was aware before this and this was the reason that I left Facebook and I sent an open letter to Zuckerberg.'
	caught off guard	surprised by info collection	I'd expect them to collect this, except the Bluetooth signals and information about Wi-Fi access points.'
		feeling unhappy	It makes you feel kind of silly.' It also makes my concerns bigger'
		fooled	That you're trying to protect yourself, but you cannot because you're using the medium and you said yes, to everything.'
	Passively discontent	feels unable to improve the situation	It's not really like I have a choice in the the matter if I want to use their application.' I think the only remedy is to call it quits'
		continued behavior	I don't want to stop using the social mediums that I'm using.' It's hard to answer. Maybe, maybe not, maybe not.' So I guess it would make me more concerned about privacy, but I don't think I act any differently.'
Privacy in Social Media	Against data collection	avoiding SM for privacy	And and and yeah, it it it was for me the reason to to get to to say goodbye to Facebook.' So that kept me from ever going on Facebook, and sometimes I do miss out on things, on kind on certain discussion groups that I know exist on Facebook.'
		disliking unknown data use	So I guess it's an unknown unknown. I I don't know what they would do with it. And I don't like that they apparently have use case for it.' Especially not when I don't know that they know.'
		scared of data use for influencing	So they determine what you what you see.'

		realizing non-privacy	I'm aware of the fact that especially using social media, uh Google or other uh, big tech firms, they know a lot about me.'
	counter code to ↑	cheap alternative	I didn't have to pay for messages like SMS.'
Social Media	information gathering	discussions	Prevent the discussions from from being very negative' So sometimes reflections on what happens in in the news.'
		feedback	It gives me a platform to to come up with my ideas and to be so thought leader' I also like it to test ideas and and write sometimes longer pieces, but also shorter pieces and get feedback on that.'
		information gathering	Yeah, maybe when you're interested in somebody, like just checking in.' I notice that you can could easily get information about news items.' see how other photographers have different perspectives on subjects' That's how I I get my information about these clubs that I'm in.'
	Social unsafety	fake news/misinformation	but it's looks so real that you think, wow, it's. Did somebody somebody really say that? And no, it's not true at all.' Hoaxes that are being talked about as if they were real.'
		real users/ genuine/commercial	it seems to be very full and also full of advertising and that is less interesting for me because then I don't feel it's very genuine what people are posting.' If I didn't meet them or have a conversation with them, I don't accept those'
		scared of other users	Use social media to spread their opinions and not only not always very nuanced'

		What I read and then I think, Oh my God, I didn't know people are … writing about that and talking about things like that in in that way'
	scamming	And you get a lot of, uh, phishing activity on Twitter, so I mean, I I'd be somewhat concerned to get hacked in some way.'
	negative comments	Prevent the discussions from being very negative and have the trolls and kind of stuff' It has become really sort of thrash thrash box.'
downside to social media	pressure	It might also make you sort of conscious that you're not working as hard as somebody else seems to be doing.' So you expect life to be good, but it's not really that good.'
	FOMO/large user base	The fear of missing out on interesting information I can use in my work' I think it's also a situation of everyone is in WhatsApp. So you start using it. For instance, if a lot of people would go away from Instagram because they were concerned, then I might also leave because then it's not as interesting anymore. Like the the typical FOMO part people have
	Keeping up appearances	Portray a better picture of what people are doing than that is actually true. It likes catch your vision of the world to bit twisted.
	negative self-image	Well, sometimes you when you're maybe not feeling too well about yourself. You think, oh, all these other people are to are working so hard and making all these beautiful things. And you constantly seeing people show the best part of the life without the bad part.
	restlessness	Social media use can create a restlessness'
leisure	fun/curiosity	Look like fun making pictures and seeing other people's pictures.

		it's still nice to see what they do and benefits is sending memes to other people, which is what I use it for make jokes with with friends And eventually I decided that that I would like to see what the people were up to. I love photography so I want to share pictures but also love love to see pictures
	relaxing	'After a day of hard work and making contracts I like some brainless activity, relaxing.'
	passing time	I was a bit bored. So I uh, I made a Facebook account. Killing time.'
	making/consuming content	there was one picture of a Kingfisher that I wanted to make for a few years and then after a few years I I had a great picture of the of the Kingfisher. It was really exciting to have that picture to be able to publish it and it I think it was really cool. I can see what other people are making and I can show other people what I'm making and that way we can connect and talk about that in comments And between actual meetings, we can still show each other what we're making.
	inspiration	And inspire each other, yeah. people that do interesting stuff and publish it on LinkedIn, that's that's a great input for me
	SM for archive	I consider Instagram as my picture diary. So sometimes: when did I do that again and I look in the pictures and then I know what year it was.' for tracking books because, (I read)
caught in the app	caught in the app	Sometimes you sort of realize that you've been longer on the app than than you intended to be, so that. planning the trip to London. I just had at that time kind of takes you out of the moment because now so when you're busy with other things in your head than just enjoying.

		it's very distracting. So it might keep you from doing your own work.
	algorithm disliking	And you notice you're in a bubble and don't get the whole picture about some news items.
	corporate power	De-platforming would be a bit of a concern. That's Instagram to lock your accounts and never give you access to it without you being able to do anything.
professionality	maintaining a large network	benefits of being on Twitter and Forecaster is being up to date with new developments in my area and making a brand for yourself so that people follow you and then you can link to your work or other presentations to get some more recognition from that. I can connect in real life with people that I know on Instagram or I can connect on Instagram with people I know in real life. Checking in like what's going on in their life. It's also my personal who's who book for people that I do not interact with on a on a regular basis, I I can get can get them back in my in my in my memory.
	image preservation	because of my role I need to be an authority and and the authority is partly. because I'm more trying to portray myself as a professional human being.
	main occupation	So you stayed there for a short time and this was a great way to stay in touch with people I did that because I was starting out with my own company
	new connections	Based on these public statements that that I make and and analysis that I make and and people responding to that and that also leads to invitations for certain conferences and and talks and so on I guess when somebody is like a headhunter or something else for a company that could be easy, but it's not what it was for me

	organizing social life	communication with friends/family	we have a a family group on Facebook. And we'll keep in contact with each other to connect with family, I guess to exchange messages and also with friends usual stuff when when people are having their birthday and so on. So so in intergroup communication, I think and sometimes a few pictures
		planning	We managed to organize uh, reunion, uh, a few years ago planned dates for face to face. So for usually for lunches and and dinners and and and and and parties and so on
Privacy	Relativity	not absolute privacy	I think privacy is OK, so I'm you know that I'm the I'm the privacy supervisor, but I think I think people go crazy a bit about privacy I'm not really concerned about my name address or data like that, but my concerns are especially, uh, in in the part of where they use my, uh, sensitive data…which can be used for identity fraud. (financial and health Sometimes I do have a photo of a flag, maybe that we have on the House and then somebody that knows Leiden might know where I live.
		liking the information collection	for example the Google Maps tracking thing I have that turned on because I sometimes I I like it and I'll go back to see where it was.
		risk/benefit	Concerned about the fact that they know where I, uh, buy my clothes or shoes or uh, that I'm a client of the Albert Heijn they weigh heavier for me than what I can gain from it or what I think I can gain from it I know the the the privacy concerns of, of course, but there's also a a balance and I at this time, I think the the the pros are are bigger than the the the cons.

	paradox	I I never really search for how to increase my privacy. For example, even though that would be a good idea How I actually use it I don't value (privacy).
Opt-in privacy	won't chase privacy	I could imagine that there is like a hidden setting somewhere that gives you extra privacy, and if I would know about the setting then I would turn it on, but otherwise I wouldn't actively search for it myself.
	choosing privacy, if it were easy	I do value privacy and whenever it's easily achievable, I try to get the most private privacy out of this.
Control	privacy as information control	I have the power myself to say whether someone can use my data Things that I rather keep to myself and not have other people know. I think privacy for me is being able to decide who you share certain information with and that it stays there That I'm the boss of who knows things about me and what people know about me. Being able to decide where and what is known about me.
	privacy measures/ (unconscious behavior)	, that's not relevant for the for the, for the larger audience or for the world. So that I don't want to share that. So there I have extra wallets and I try to obfuscate that's a bit so that people can see it. It's a habit it's not really an action, it's just.
Privacy activist	scared when too much info is public	if everything were known about me and uh, publicly available, I would feel a bit more vulnerable, I think.
	location privacy	I do tell say that I live in Leiden, but I don't show my front door or my house number or the view for my car. So things about where I live and all these kinds of things, I I don't want to. I don't. I don't want to. To

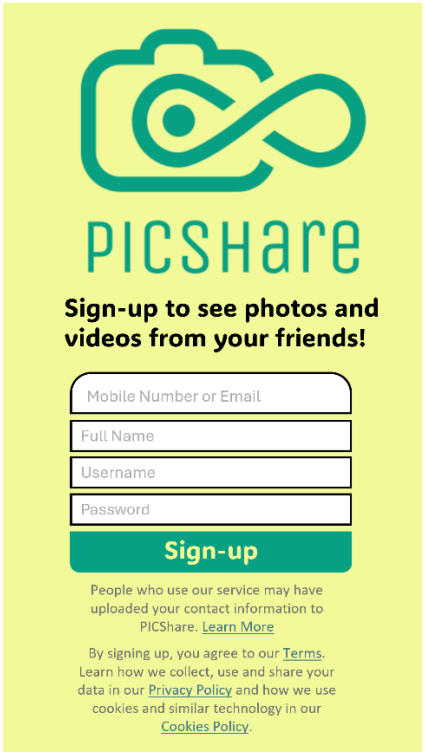
		to mention too much or so some some people know of course, but. So I don't, for instance, give the location of where I take a photo.
	Concerned for privacy of family	Things I say to perhaps my girlfriend or my my, my, my mother, but I would like to keep private for myself. I also don't link to my children to. And so I have a family, but I don't. I don't necessarily publish about them or post pictures about them. So so that that is for them to decide
	liking privacy/privacy activist	I think privacy is very important. I sent an open letter to to to Zuckerberg

Appendix E: Survey questions overview

Description of PicShare

For this study, please consider this fictional example and answer the following question:
PicShare is social media app/service that allows you to share photos and videos with your friends. Almost all people from your generation use PicShare. It is used for memes, chatting, uploading and watching pictures and videos of you, your friends, family, celebrities and influencers.

Control group



Treatment group

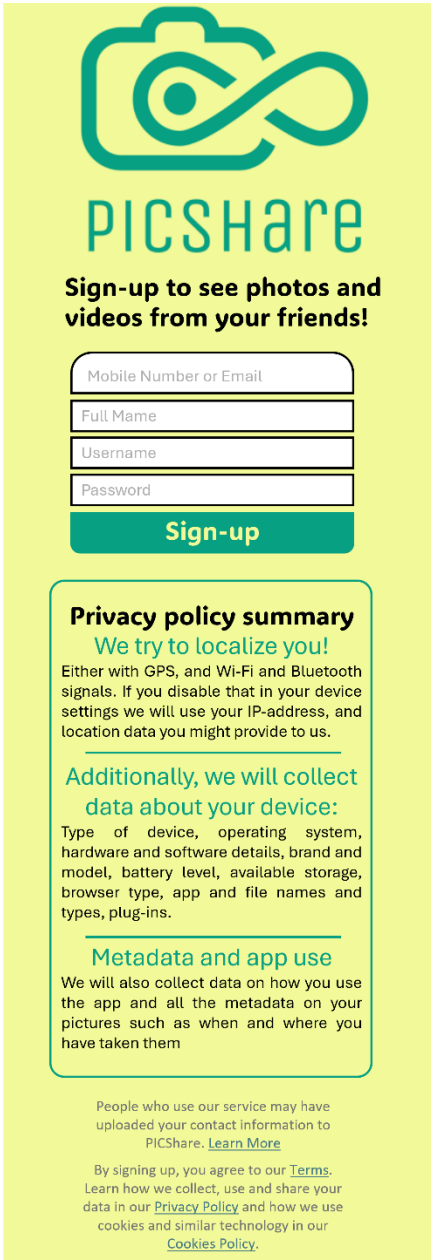


Table 7

Overview of the survey questions, answering options and source or purpose.

category (nr. of questions)	question code	Question	Answer type	Source/purpose
experiment (1)	1	I would make an account on this service	7-point Likert scale*	main experiment
Performance (5)		Using social media can enable me to:		
	P1	plan something with friends and family	7-point Likert scale*	Balakrishnan, 2016, quantitative research
	P2	relax and have a laugh	7-point Likert scale*	Balakrishnan, 2016, quantitative research
	P3	get information on people and events	7-point Likert scale*	Balakrishnan, 2016, quantitative research
	P4	keep in touch with professional contacts	7-point Likert scale*	Balakrishnan, 2016, quantitative research
	P5	keep in touch with friends and family	7-point Likert scale*	Balakrishnan, 2016, quantitative research
Self (5)	S1	I can learn easily to use social media	7-point Likert scale*	Balakrishnan, 2016
	S2	I have skills to use social media	7-point Likert scale*	Balakrishnan, 2016
	S3	I have confidence to use social media	7-point Likert scale*	Balakrishnan, 2016
	S4	I take a positive attitude towards myself when I use social media	7-point Likert scale*	Balakrishnan, 2016
	S5	It enhances my interpersonal relationships with others	7-point Likert scale*	Balakrishnan, 2016
Communication functionality (5)		Using social media, I can:		
	CF1	Communicate comfortably with peers (compare to face-to-face communication)	7-point Likert scale*	Balakrishnan, 2016
	CF2	Communicate with peers more conveniently	7-point Likert scale*	Balakrishnan, 2016
	CF3	Improve communication skills with others	7-point Likert scale*	Balakrishnan, 2016
	CF4	Express myself to others.	7-point Likert scale*	
	CF5	Share announcements/news matters easily	7-point Likert scale*	Balakrishnan, 2016

Influence/effort (5)		I use social media because:		
	IE1	It is popular	7-point Likert scale*	Balakrishnan, 2016, quantitative research
	IE2	My friends are using it	7-point Likert scale*	Balakrishnan, 2016, quantitative research
	IE3	Interacting with social media is clear and understandable	7-point Likert scale*	Balakrishnan, 2016
	IE4	Learning how to use social media is easy	7-point Likert scale*	Balakrishnan, 2016
	IE5	It is easy to use	7-point Likert scale*	Balakrishnan, 2016
Privacy awareness and preferences (5)	PAP1	I know all the parties who collect the information I provide during the use of the app.	7-point Likert scale*	Choi et al., 2023
	PAP2	I am aware of the exact nature of information that is collected during the use of the app.	7-point Likert scale*	Choi et al., 2023
	PAP3	I am not concerned that the information I submitted on the app could be misused.	7-point Likert scale*	Choi et al., 2023
	PAP4	I want to be in control of what data of myself is being used	7-point Likert scale*	qualitative research
	PAP5	I care about my privacy	7-point Likert scale*	qualitative research
Characteristics (7)		What is your gender?	Male, female third/no gender	Venkatesh et al., 2003/for diversification of respondents
		How old are you?	numerical	Venkatesh et al., 2003/for diversification of respondents
		How often do you use social media?	Every Hour, Most Hours, Every Day, Most Days, Once a week, less often, Never	Venkatesh et al., 2003/for diversification of respondents
		what is your current or highest education?	High School, practical/mbo, applied higher education/HBO, theoretical higher education/WO, Master, PhD, none	for diversification of respondents
		What is your place of residence?	text answer	for diversification of respondents
		What types of social media do you use?	Multiple answers: Instagram, Facebook, Snapchat, TikTok, LinkedIn, WhatsApp, Other, namely:, None	for diversification of respondents/validation of answers
		I decide whether I use social media.	7-point Likert scale*	Venkatesh et al., 2003/for diversification of respondents

Note: *Strongly Disagree, Disagree, Somewhat Agree, Neither Agree, Nor Disagree, Somewhat Agree, Agree, Strongly Agree

Appendix F: Summary Statistics

Table 8

Summary statistics of the control variables

Category	Frequency	Percentage
High School	2	2,3
Practical/MBO	1	1,1
Applied Higher Education/HBO	14	16,1
Theoretical Higher Education/WO	41	47,1
Master	29	33,3
Every hour	11	12,6
Most hours	34	39,1
Every day	37	42,5
Most days	4	4,6
Once a week	1	1,1
Instagram	83	95,4
WhatsApp	86	98,9
snapchat	74	85,1
LinkedIn	73	83,9
TikTok	35	40,2
Facebook	33	37,9
twitter/X	10	11,5
other	10	11,5
Zuid-Holland	78	89,7
Utrecht	3	3,4
Zeeland	2	2,3
Netherlands General	4	4,6
I decide whether I use social media		
Strongly disagree	1	1,1
Disagree	3	3,4
Somewhat disagree	13	14,9
Neither agree nor disagree	5	5,7
Somewhat agree	29	33,3
Agree	22	25,3
Strongly agree	14	16,1

Table 9

Summary statistics of the determinants

	Mean	Median	Std. Dev.
P1	5,29	6	1,67
P2	5,85	6	0,934
P3	6,02	6	0,821
P4	4,6	5	1,551
P5	5,84	6	1,15
S1	6,08	6	0,852
S2	5,75	6	1,112
S3	5,68	6	1,062
S4	4,66	5	1,199
S5	4,83	5	1,059
CF1	4,46	5	1,274
CF2	4,77	5	1,412
CF3	3,62	3	1,416
CF4	4,38	5	1,4
CF5	5,75	6	0,93
IE1	4,7	5	1,699
IE2	5,62	6	1,305
IE3	4,74	5	1,325
IE4	5,11	6	1,368
IE5	5,59	6	1,063
PAP1	1,99	2	1,196
PAP2	2,28	2	1,395
PAP3	3,28	3	1,476
PAP4	5,08	5	1,314
PAP5	5,45	5	1,02

Appendix G: Quantitative Results of sub question 4 & 5

Table 10

Factor Analysis

	Factors								
	NC	EU	TWC	FE	Pop	PA	PP	SI	OWC
P4	0,726								
P1	0,711								
S5	-								
P5	-								
IE4		0,92							
IE5		0,912							
IE3		0,76							
CF3			0,76						
CF1			0,746						
CF2			0,734						
CF4			-						
S2				0,879					
S3				0,779					
S1				0,725					
IE2					0,875				
IE1					0,861				
PAP1						0,922			
PAP2						0,91			
PAP4							0,752		
PAP5							0,712		
PAP3							-		
S4								0,796	
P3								-	
P2								-	
CF5									0,863

Note: [a] Extraction Method: Principal Component Analysis, [b] Rotation Method: Varimax with Kaiser Normalization, [c] the rotation converged in 12 iterations.

Table 11

Independent sample t-test of experiment on Privacy Awareness and Preference questions

	Mean control N=43	Mean treatment N=44	Mean difference
PAP1	1,91	2,07	-0,16
PAP2	2,14	2,41	-0,27
PAP3	3,28	3,27	0,01
PAP4	5,12	5,05	0,07
PAP5	5,53	5,36	0,17

Note: * p < ,1, ** p < ,05 *** p<,001.

Appendix H: Linear regression assumptions.

Linearity: All the scatterplots looked somewhat linear.

Figure 2

Experiment question plotted against TWC

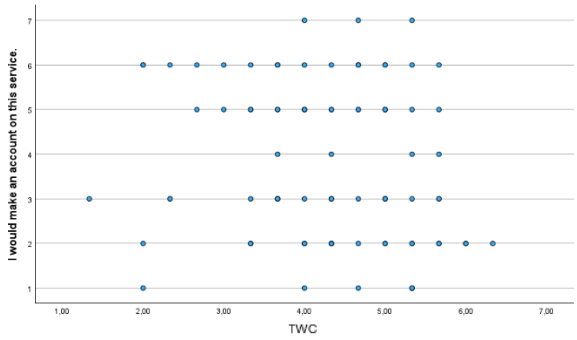


Figure 3

Experiment question plotted against FE

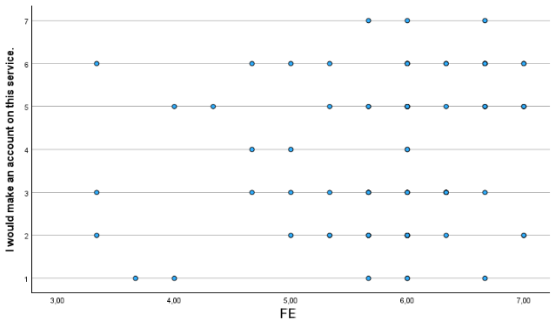


Figure 4

Experiment question plotted against Pop

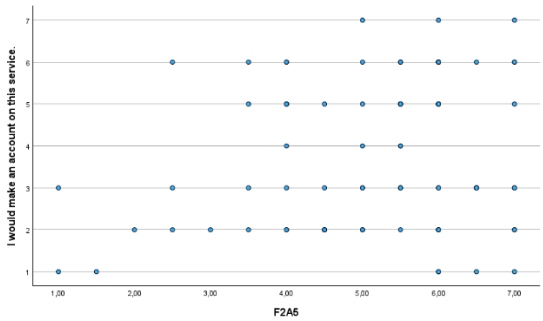
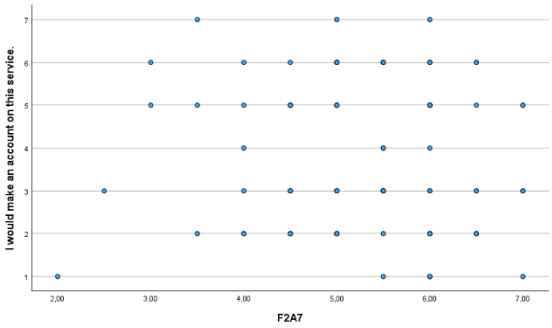


Figure 5

Experiment question plotted against PP



Independence: The results seem to be independent, because the Durbin-Watson test resulted in 2,253 for model 6 and 2,163 for model 8, both are between 1,5 and 2,5.

Homoskedacity: The unstandardized residuals seem evenly distributed when plotted against the unstandardized predicted values, so homoskedacity can be assumed.

Figure 6

Unstandardized Residuals plotted against Unstandardized predicted value of model 6

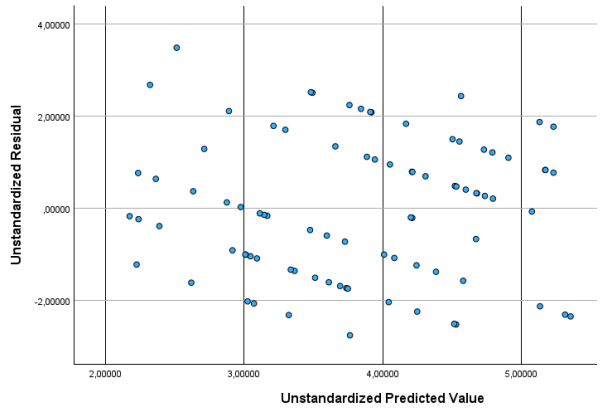
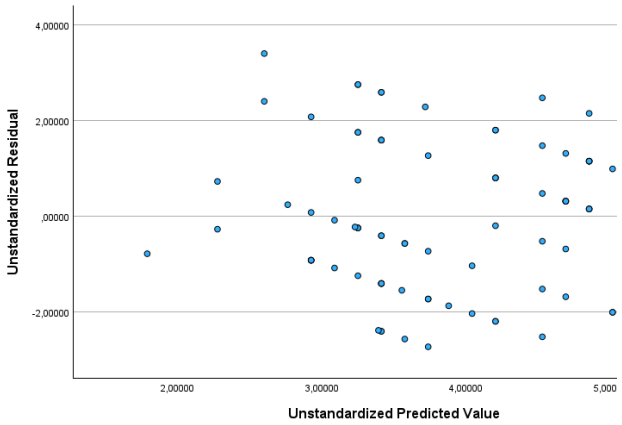


Figure 7

Unstandardized Residuals plotted against Unstandardized predicted value of model 8



Normal distribution of residuals:

Shapiro-Wilks had a significance level of 0,109 for model 6 and a significance level of 0,06 for model 8, so the hypothesis of normal distribution cannot be rejected. However, when looking at the histograms there is a skewedness to the left meaning that the residuals are containing information that negatively impacts the prediction of IMSMA.

Figure 8

Histogram of the Unstandardized Residuals of model 6

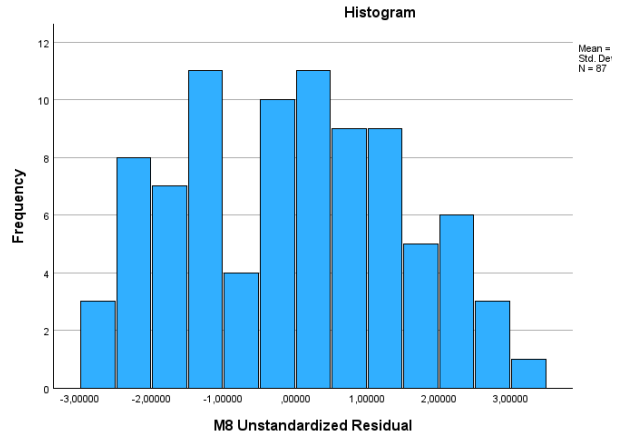
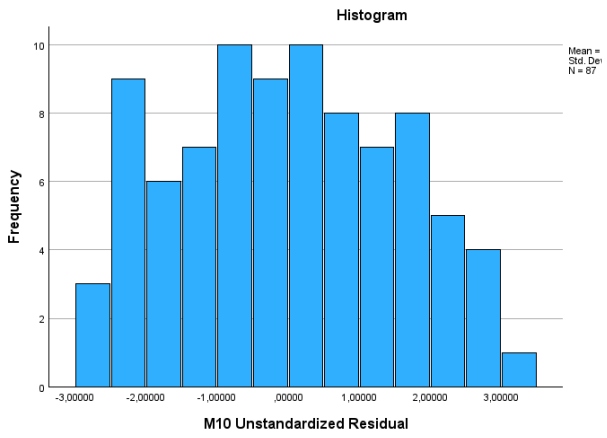


Figure 9

Histogram of the Unstandardized Residuals of model 8



Multicollinearity:

Multicollinearity might exist between Popularity and Privacy Preferences in model 6, based on the correlation matrix. The correlation matrix of model 8 gives no indication of multicollinearity.

Table 12

Correlation matrix of model 6

		TWC	FE	Pop	PP	Dummy Experiment
TWC	Pearson Corr.	1	-,018	,108	,019	,168
	Sig. (2-tailed)		,871	,319	,864	,120
FE	Pearson Corr.	-,018	1	,082	-,012	-,048
	Sig. (2-tailed)	,871		,448	,915	,657
Pop	Pearson Corr.	,108	,082	1	,269*	,148
	Sig. (2-tailed)	,319	,448		,012	,172
PP	Pearson Corr.	,019	-,012	,269*	1	-,058
	Sig. (2-tailed)	,864	,915	,012		,592
Dummy Experiment	Pearson Corr.	,168	-,048	,148	-,058	1
	Sig. (2-tailed)	,120	,657	,172	,592	

Note: * $p < .05$, [a] $N = 87$

Table 13

Correlation matrix of model 8

		Pop	Dummy Experiment
Pop	Pearson Corr.	1	,148
	Sig. (2-tailed)		,172
Dummy Experiment	Pearson Corr.	,148	1
	Sig. (2-tailed)	,172	

Note: * $p < .05$, [a] $N = 87$