Social Privacy on Facebook:
A cross-sectional survey analyzing awareness among university students in the Netherlands

Student Name: Metin Yazici
Student Number: 455219
Supervisor: Ju-Sung (Jay) Lee, PhD

Media, Culture & Society
Erasmus School of History, Culture and Communication
Erasmus University Rotterdam

Master Thesis
July 2017

Social Privacy on Facebook: A cross-sectional survey analyzing awareness among university students in the Netherlands

ABSTRACT

Social privacy originates from the user's knowledge and strategies to control their personal information shared on social network sites. Social network sites have changed the patterns of disclosure and dissemination practices of personal information. Being aware of social privacy is a condition that has became significant in our lives, along with the ubiquitous presence of social network sites. Social privacy consists of any set of circumstances involving the control of personal information disclosed on social network sites over exercises of surveillance happening between individuals. Recent studies illustrate that social privacy is a great concern for the majority of social network site users. The aim of this thesis is to examine the extent of social privacy awareness among university students in the Netherlands, along with the questions of what students disclose on their Facebook profiles and to which audience they disclose. The use of personal information disclosure and visibility strategies has been examined in detail using a quantitative approach. A cross-sectional survey is conducted through random sampling of the students studying in the Netherlands ($N = 176$). The results showed that social privacy awareness has a strong association with negative social network site experiences and a less strong association with the use of technological privacy tools and intensity of Facebook use. Contrary to general expectations, undesired visibility and surveillance are a great privacy concern for most students. After all, the university students in the Netherlands are aware of social privacy to various extents, depending on other variables; the associations with default privacy settings and negative experiences are found to be substantial. Future research is recommended to assess the in-depth relationship between personal information disclosure and social surveillance by way of individualized items and consistent questions. This will need to examine information sharing motivations and the concerns of surveillance practices.

Keywords: *social, privacy, surveillance, disclosure, visibility*

PREFACE

I deeply thank my supervisor, Ju-Sung (Jay) Lee, PhD, who lit the path in my journey towards my first comprehensive academic writing by sharing his invaluable knowledge and experience. I would like to thank my family for their everlasting support in making this master possible. I would also like to thank my dear Dominika for believing in me all the time.

Table of Contents

List of Tables

List of Figures

1 Introduction

Social network sites[1], or social media, no longer belong to an exclusive sphere that once was available only to a specific group or community; instead, it is a ubiquitous and vibrant place for almost everybody, enabling users to portray their own identity and interact with each other. They have changed the patterns of personal information disclosure and dissemination practices of personal information, by allowing everyone to share their lives online. They have the ability to reveal personal information as sorts of "network[-ing] softwares" (Trottier, 2010) which make user information accessible, not bound by time and space limitations. That is to say, the information disclosed on social network sites exceeds the privacy boundaries because the management of personal information on social network sites requires a different approach and strategies to preserve it.

The significance of boundaries of personal information has declined in the course of the development of social network sites. In physical life, people disclose their thoughts and actions in a limited time and space network, whereas online life offers a network without this limit, in which time and space limitations of personal information are everchanging. The increment in social and communicative acts staying permanent in the digital world is risky: "...what was once ephemeral, with evidence of it living only in the memory of the current witnesses (Tufekci, 2008, p. 21)." because of the reason that the information on social network site has backward capacity which turns it something traceable. As an example of this, in a conversation in the street, an opinion expressed or a greeting with a person in a place, though it leaves behind an impression on both participants, they might not be able to even remember what was said or seen in the future. The change of communication network practices in the digital world has created traceable and reproducible data such that the activities surrounding use of these data has raised concerns over privacy.

The issue of protecting privacy has a long history even before the advent of the Internet. The notion of privacy is frequently affiliated with individual privacy which has always been preserved — it has gained its legitimacy in Western understanding throughout 200 years of liberal ideas about individual rights in personal information (Nissenbaum, 2009). Privacy is a socially-constructed concept that shows a clear privacy definition is not valid. Nevertheless, a

[1] Boyd and Ellison (2008, p. 211) emphasize that "network" is more apposite than "networking" for describing this phenomenon. While networking implies a relationship often happening between strangers, network emphasizes that they already live in a network.

clear privacy definition may be required because privacy is a phenomenon existing in daily life. Privacy is an interest and management of how much information is able to be maintained by individuals (Sloan & Warner, 2013). Users should have control over their personal information and who can access it (O'Brien & Torres, 2012).

In legal aspects, personal information typically is considered to be personally identifiable (or identifying) information that can be used to identify a single person in some aspects. Personal information has been transformed with the advent of social network sites; however, it has never been totally private or intimate. Moreover, personal differs from private information that generally appears in private messaging services, i.e. content sent to some friends considered private information rather than personal information. The "privateness" of the messaging platform itself indicates the degree of privacy of the information, no matter what the content is or how many people receive it. Private information is not in the scope of this thesis research which does not show characteristics of social privacy and surveillance.

One of the most persuasive motivations for why users take advantage of social network sites and prefer to disclose information is that they want to be seen (Tufekci, 2008). Producing and disclosing personal information online have become essential attributes in the contemporary world. According to Eurobarometer (2015), a "large majority of people (71%) still say that providing personal information is an increasing part of modern life and accept that there is no alternative." It is no surprise that people prefer to disclose information on social network sites in order to be satisfied and avoid exclusion from the social sphere (Debatin, Lovejoy, Horn, & Hughes, 2009). Owing to a growing popularity of social network sites, sharing personal information becomes the new desirable thing. Nevertheless, setting privacy boundaries is desired by people who do not want every piece of personal information, which basically identifies the individual, to be visible to all.

Social privacy is about control of personal information data from unwanted visibility. The unwanted visibility also mean unwanted audiences that are those to whom the user did not want their personal data to be disclosed. The disclosure of personal data can be achieved by maintaining any amount of personal information or access to other individuals. Therefore, users want to manage who is going to *surveil* them when disclosing personal information to the public is somehow inevitable. Social privacy is perceived as a protection against surveillance practices sourced from disclosure of personal information. Surveillance in social network sites has transformed the maintenance of interpersonal relationships (Tokunaga, 2016). Users can have greater concern of the privacy intrusions coming from other persons' surveillance, which is how their friends, family, or acquaintances see their personal information (Young & Quan-Haase, 2013). On that matter, the concept of social privacy should be considered alongside social surveillance.

Social surveillance (also known as "interpersonal surveillance" (Trottier, 2012a) or "interpersonal electronic surveillance" (Tokunaga, 2011)) is a term depicted by (Marwick, 2012,

p. 382), and is an “ongoing eavesdropping, investigation, gossip and inquiry that constitutes information gathering by people about their peers...” Social surveillance is one of the results of having a profile on the social network sites. Trottier (2012a) identifies that users are aware of surveillance practices on Facebook, a platform which is already designed to share personal information online. Users can continually investigate how friends, family, and acquaintances are living, what they think, and what they are doing owing to social network sites. Since motivations of disclosure consist of the wants of attention, social status, and visibility of personal details to an audience, social surveillance should be considered as “partly consensual” (Marwick, 2012).

Social network sites make surveillance more apparent in society as personal information on these sites alter the physical boundaries and make it accessible in the Internet network (Trottier, 2012b). For instance, someone declaring that they took drugs or have alcohol problems might make a negative impression on others. In addition to that, employers use social network sites to monitor personal information with the intent of screening their applicants and surveilling their employees. It is reported that 60% of employers use social network sites to research candidates and more than 25% of employers have stated that their content online has induced a reprimanding or firing (Chad, 2016). To that respect, the understanding of social privacy is subject to the changes in the level of information disclosure and social surveillance.

Facebook, which is a multi-layered platform possessing the characteristics of user-generated content, defines itself as empowering people to share and make the world more connected (Facebook, 2016). Users are able to share their personal material on Facebook, such as their name, age, date of birth, profile picture, photos, videos, relationship status, and other kinds of information which does not usually “offend the eye” (e.g., sharing naked/sexual photos on a profile may not be suited to the intention of Facebook. Facebook is a fruitful place to perform surveillant practices over individuals because the platform motivates users to create content mainly produced by personal information. Facebook users have a variety of tools to control their personal information. Nigam (2013) gives some instances regarding how users can increase their social privacy online: users can limit the posts from “who can see”, arranging to “friends only” under the privacy tools; refrain from using “public” settings when sharing content; accept friend requests only from known individual users; cautiously control degree of accessibility through search engines and have power over who is able to view the personal information; facilitate the “post review” action helping control tagging of photos or other kinds of related content; and review their own profiles habitually and delete or hide undesirable content from the News Feed.

Facebook is the first and longest-lasting, though not the oldest, social network site to gain a rapid popularity over the world in a short-period of time. Facebook is a valid platform to look for users appertaining to different socio-demographic characteristics, due to its widespread usage. Facebook has been a popular topic in the last years due to its great popularity and diffusion over platforms. According to Newcom (2017), Facebook is the second-most popular social network site in the Netherlands (just after Whatsapp — which can be considered a communication platform rather than a social network site). 10.4 million people use Facebook with a 7.5 million daily penetration. Facebook is very popular among 16-18 year olds (80%) and 19-25 year olds (89%)[2].

The available research in online privacy is various. Achieving online privacy through privacy tools, like settings, on social network sites has been researched by many scholars (e.g., Debatin et al., 2009; Lewis, Kaufman, and Christakis, 2008; Waters and Ackerman, 2011) including its benefits and risks (e.g., Andrejevic, 2005; Joinson and Paine, 2007; O'Brien and Torres, 2012). Privacy on social network sites is mainly examined specifically related to disclosure practices (e.g., Acquisti and Gross, 2006; O'Brien and Torres, 2012; Tufekci, 2008). Social privacy is a superior concern over institutional privacy because users are more concerned about the ways they are exposed to the people they know rather than how their data is obtained by corporations and governments (Boyd & Hargittai, 2010). Some researchers (e.g., Acquisti and Gross, 2006; Boyd and Hargittai, 2010; Young and Quan-Haase, 2013) have focused on both social and institutional privacy and disclosure; however, there is limited work on the social aspects of privacy in relation to social surveillance and information disclosure. Even so, they do not approach the problem the same way this thesis does, in terms of explaining social privacy. This thesis approaches the social privacy issue on social network sites along with social surveillance and information disclosure practices, which are preserved through visibility strategies that are basically based on the number of ways to keep unwanted audiences away from personal information.

This thesis focuses on university students' awareness of social privacy on Facebook to explore the kind and amount of information users disclose, the strategies they employ, their concerns about other users' gazes, and how they make that information accessible to others. Being aware of social privacy is based on users' awareness of controlling visibility and personal information disclosure practices. This study conceives the concept of social privacy as a phenomenon existing between surveillance and disclosure practices. Thereby, the thesis wishes to fulfill the question:

> ***RQ:*** *To what extent are university students in the Netherlands aware of social privacy on Facebook?*

Academic Relevance

This thesis endeavors to fill a gap in the literature regarding awareness of social privacy, which

[2] There is a segregation between social network sites on their usage such as Twitter for political participation and news media (Kwak, Lee, Park, & Moon, 2010), Instagram for self-presentation through photo sharing, or LinkedIn for professional work-related content (Bakhshi, Shamma, & Gilbert, 2014). Facebook is much broader and comprises multi-layered features that are answering the user's needs, generally, "for everything". Moreover, Facebook is the most populated social network platform.

is in relation to social surveillance and information disclosure practices along with the visibility strategies of social network sites, notably Facebook. The research question is helpful to answer “what is shared” in the context of disclosure and “to whom is shared” by means of the practices/strategies of visibility. The approach of this thesis is different than the previous research which perceives social privacy awareness in corporation with social surveillance and visibility strategies.

The survey research aims to acquire more generalizable results about students’ social privacy awareness (Boyd & Marwick, 2011). Students in Dutch universities are the central subjects of the research. A cross-sectional survey has been implemented on the unit of analysis using university students in the Netherlands who have a Facebook account or have had one in the past. This thesis strives to explore the general influential factors of the subject matter in a bigger picture; therefore, a survey has the potential to capture more general features that qualitative analysis, such as interviews, may not capture due to narrow capability. Mostly, Facebook is the first and only place for many students — especially in the Western world — to disclose personal information in order to interact with friends socially, build networks, and even form identity. Nevertheless, Facebook has not only changed information and communication practices but also surveillance practices of disclosed information. The main objective is to reveal the insights into what extent social privacy awareness is present among student users on Facebook, and also show quantitative correlations with variables.

Societal Relevance

Studying awareness of social privacy on social network sites is significant to illustrating societal behaviors and structures, along with the preservationist strategies regarding this issue that would contribute to existing academic literature. The attention given to this issue is not more than a mere consideration of privacy in all dimensions, which surrounds the majority of existing privacy studies. Previous studies have researched privacy awareness in relation to users’ patterns and behavior on social network sites (e.g. Acquisti and Gross, 2006; Govani and Pashley, 2005; Tuunainen, Pitkänen, and Hovi, 2009).

The majority of university students use social network sites for sharing information and developing networks (Raacke & Bonds-Raacke, 2008). Contrary to general belief, undesired visibility and surveillance is a great privacy concern for most people. The level of confidence issues in privacy on social media is not low in the Netherlands (Newcom, 2017). Fifty-four percent of users are worried about data provided online, but only 17% have complete confidence in social media. Young people have more concern about their privacy and are aware of the potential dangers of low privacy, taking precautions manifested through identities, attitudes, and behaviors (Boyd & Marwick, 2011; Debatin et al., 2009). Based on this premise, the current students largely belong to the group called “Millenials” (e.g., Livingstone, 2013; Waters and Ackerman, 2011), a generation born between the 1980s and 1990s, and “Digital

Natives[3]" (e.g., Kaplan and Haenlein, 2010; Williams, Crittenden, Keo, and McCarty, 2012), a generation born between the 1980s and 2000s, whose engagement with the digital world is surely different and worth investigating.

Privacy is a vital issue for society and should be debated in each and every time and space, and also should be given particular attention at times when technology takes a sharp leap. Regan (1995) emphasizes that society functions "healthier" when privacy exists. The state of privacy, which adjusts based on "who has discretion or control over determining the degree of access (Nissenbaum, 2009, p. 70)" between individuals in society, is a stationary condition as long as society and social interaction exist together. Since social network sites are a public sphere embracing the regulation of the private lives of individuals, studying moral, behavioral, and societal actions of privacy is significant. The state of privacy should be widely projected as a whole instead of merely thinking of individual privacy interests. All in all, the presence of "online" takes the privacy issue further and to a broader platform, which makes privacy more significant than ever before.

Thesis Outline

This thesis has the following organization. The Theoretical Framework chapter reviews the existing literature on social privacy, social surveillance, information disclosure, technological privacy tools, and previous negative social media experiences and forms the hypotheses. The Method chapter follows research design, population and sampling, and data collection and analysis, and operationalized variables are explained in detail. The Results chapter presents the descriptive and inferential statistical interpretations of the survey data. Therefore, the discussion of the results were integrated to this chapter. Lastly, the Conclusion chapter finalizes this thesis with ending remarks from the findings, calling back the literature, discussing the limitations, and proposing suggestions for possible future research.

[3] Digital Natives, named by Prensky (2001), implies the generation who are native speakers of digital world such as computers, video games, the Internet, etc.

2 Theoretical Framework

2.1 Privacy on Social Network Sites

Privacy is a socially-constructed concept incorporating the norms and values of everyday life, so the universal definition of privacy as being valid and trusted all the time is not quite possible (Boyd & Marwick, 2011). The notion of privacy cannot essentially hold any universal explanation and remain unaffected through different times and spaces. Instead, privacy dwells in a constant state of flux, considering that subjective human interaction is objectified by the institutional world, which is man-made and constructed objectivity (Berger & Luckmann, 1991). Even so, the existing literature about the notion of privacy has shown that there is not a common agreement on its definition. The oldest and sufficiently popular definition of privacy is "the right to be let alone" (Warren & Brandeis, 1890) and to prevent all intrusions from harming personal freedom and dignity (Bloustein, 1964). There is not one essential existence of privacy outside of social situations in human relationships (Solove, 2002); privacy is neither a series of universally accepted human activities, nor is it an independent natural element or part of reality (Gutwirth, 2002, p. 29), and it is a socially-constructed norm that reflects the values and norms of individuals (Boyd & Marwick, 2011). The notion of privacy should be contextually considered, as different situations and subjects may bear different contextualisations. Every individual has different concerns about its own privacy that are managed by different perceptions.

Despite the relativity in social context, in this thesis the concept of privacy is involved with the management of how much to disclose of oneself and how much share to share with others (Altman, 1975; as cited in Tufekci, 2008). Privacy is conceived as an interest, whereby individuals are able to maintain the information they provide (Sloan & Warner, 2013), and in addition avoid interference by others (Clarke, 1999). Every user should have control over the personal information they disclose and who is able to access it (O'Brien & Torres, 2012).

2.1.1 Social Network Sites and Personal Information

Social network sites have most certainly changed the ways of communication, entertainment, and sources of information, owing to its self-producing and ever-growing nature (Shao, 2009).

Social media, or social network sites, is the subcategory of user-generated content[1][2] defined by Boyd and Ellison (2008, p. 211) as a "web-based service" allowing users to (I) create a public or semi-public profile in a bounded platform, (II) have the ability of "articulating a list of other users with whom they share a connection", and (III) view the user's list of connections. There are users who maintain their level of information and communication by creating new content and changing the level of user access (Blank & Reisdorf, 2012). Thereby, social network sites enable users to connect by creating personal profiles, inviting friends and colleagues to have access to those profiles, and sending e-mails and instant messages to each other that have led to a transformation from a communication culture based on consumption to a "communication culture of participation" (Wyrwoll, 2014).

2.1.2 The User in Social Network Sites

The user, who is the most prominent foundation of the production process on social network sites, can be considered under two different sub-categories: producing-user and consuming-user. A producing-user is producing the content, and a consuming-user is receiving and consuming the content produced by the producing-user in the social network sites. Nevertheless, the difference of users in the present online community era is not sharply divided; these two are merged into one category of "prosumer[3]". The emergence of the name "prosumer" is constituted by the combination of the two concepts, which are "users" who make active contributions to the Internet outside of professional routines (OECD, 2007; Ritzer, Dean, & Jurgenson, 2012; van Dijck, 2009) and "consumers" of the media (Grinnell, 2009). The functions of social network sites are heavily dependent upon the users; in most cases, sites could not exist at all without the work done by a "crowd of prosumers" (Ritzer et al., 2012).

The dimensions of disclosure are the user's production of its own personal information. The content, which is in the form of "personal information" created by users, is the fundamental material for the existence of social network sites. In principal, the thesis engages with the concept of personal information which regards the sum of all different kinds of factual data appointing a private individual user as identifiable. Facebook as a social network site is able to

[1] User-generated content has several different names in the literature such as *Web 2.0 technologies* (O'Reilly & Battelle, 2009); *Social Web* (Gruber, 2008); *social media*; *computer-mediated communication*; *consumer-generated media*; and *social network sites* (Boyd & Ellison, 2008).

[2] User-generated content has three main distinctive features (OECD, 2007), which are slightly different than the social network sites. First, the content should be publicly accessible over the Internet. Second, the produced content requires "a certain amount of creative effort." Third, user is an agent producing content, "who chooses to work within and those who choose to work outside professional routines and practices (p. 45)." The first condition excludes contents such as e-mails or instant messages; second, excluding the replications of existing content (i.e. copy and paste content, or re-tweet) (Kaplan & Haenlein, 2010); third, about the content created by non-professional users.

[3] American futurist author Alvin Toffler coins the term "prosumer" for the first time in his book *The Third Wave*, published in 1980. However, the term has been changed over years in the media studies.

operate with the existence of personal information data disclosed by users. Although the concept of prosumer is significant in today's convergent network, the separation of users between producer to consumer is needed to avoid confusion in terms. Producing-user and consuming-user are particularly construed as the "user[4]" and "audience[5]" respectively in this thesis.

2.1.3 Networked Sites, Boundaries and Privacy

Social network sites are a sort of "network[-*ing*] software" storing information in a single location (Trottier, 2010), which makes privacy hardly achievable due to "searchability", "persistence" and "cross-indexability" of personal information (Tufekci, 2008). The personal information of individual users is no longer to be discovered, rather it could be easily approachable by someone regardless of time and space (Park, Shin, & Ju, 2015). The accessibility of personal information has been altered due to the redefinition of boundaries. Personal information available online renders the scope to exceed privacy boundaries (Trottier, 2010). Individuals have materially defined boundaries in the physical world (Boyd & Ellison, 2008); for instance, an information or action disclosed in an enclosed place, like a room inside walls, could be maintained by the individual owing to the encapsulation of dimensions of time and space (except capturing elements of physical life through methods such as recording devices — i.e. photo machine, voice recorder, "writing" — may overcome this). Personal information disclosed through social network sites is not subject to time and space dimension boundaries. That is to say, boundaries are radically altered, and procedures to protect privacy become defenseless in the online world.

Boyd and Marwick (2011) mentions the term "networked publics", referring to a space where the audience can interact with each other. Networked publics is "the space constructed through networked technologies and the imagined community that emerges as a result of the intersection of people, technology, and practice (p. 7)." Facebook is a networked public that ensures the flow of information and communication and allows individuals to gather in a single place, or platform. Building a persona and establishing communication in online is so embedded with engagement and participation in the networking publics. Even so, not existing on social platforms might also mean not existing in physical life in some positions (Debatin et al., 2009). For instance, if a young student has no access to or prefers not to be exist on social media, this can lead them to having a lack of identity or communication in various communities like school, work, friends and even family. Even so, it is common that people ask the reason why a person does not have a Facebook account that having a Facebook profile might be a kind of excuse for some cases because of preferring not to exist on social media.

[4] "User" may also be occasionally used in the generic form, which implies both characteristics of the producing and consuming user.

[5] It is also called a "receiver" in some cases.

2.2 Social Privacy

Personal information was not something private or intimate at all before the advent of social network sites, as the information had been disclosed to certain known people. Social network sites transform the availability and accessibility of personal information into an indication unbounded by time and space. The content shared on social network sites pertains to the public sphere because the personal information shared with the other individuals. Personal information appears online in two ways: first, the user (or actor) generates data voluntarily; in other words, disclosing personal information, which is basically anything that users knowingly and willfully upload to social media — the kind of personal information this thesis takes as a unit. Second, user data is generated from a direct or indirect result emerging from users' tasks or behaviors (Zwitter, 2014). Social privacy is about controlling the access to personal information data from social surveillance practices[6] and attempts to preserve privacy by keeping an "unwanted audience" away from personal information. Boyd and Marwick (2011, p. 4) argues that social privacy can be achieved by establishing "access" to other individuals, and that "loss of privacy occurs as others obtain information about an individual, pay attention to him, or gain access to him."

Social privacy should be conjointly thought of with social surveillance, which brings the initial meaning to what social privacy is, initially discussed to clarify its meaning. Social surveillance has been discussed to a great extent in the literature (e.g., Albrechtslund, 2008; Andrejevic, 2005; Boyd and Marwick, 2011; Fuchs and Trottier, 2015; Marwick, 2012; Tokunaga, 2011; Trottier, 2012a), and all show the essentiality of surveillance happening within an individual-based relationship activity in social network sites.

2.2.1 Social Surveillance

Surveillance happening between individuals on social network sites is a phenomenon identified with different names in the literature; this thesis employs the term "social surveillance" to define this kind of *surveillant* practices. In simple terms, social surveillance — through social network sites — is "examining contents created by others and looking at one's own content through other people's eyes (Marwick, 2012, p. 378)". Andrejevic (2005) defines "lateral surveillance" as individuals gaining information about others, like family, relatives, spouses, friends, and colleagues (which Andrejevic (2005) defines "peer-to-peer surveillance" derived from the two examples). Albrechtslund (2008) derives the social aspect of surveillance from a concept "participatory surveillance", where individuals voluntarily disclose personal information

[6] Institutional privacy is about the concerns of how third parties use personal data (e.g. aggregating users' personal data through computational methods to use for marketing benefits) whereas social privacy relates to surveillance practices from interpersonal relations (Young & Quan-Haase, 2013). Institutional privacy is not the scope of this thesis research.

for making their activities, like preferences, tastes, opinions etc., visible. Trottier (2012b) employs the term "interpersonal surveillance", which is the most proximate term for social surveillance, highlighting types of surveillance occurring between individuals. Fuchs and Trottier (2015) identifies surveillance happening between individuals with regard to the convergence of the different social roles of individual, where surveillance becomes the "monitoring of different activities in different social roles with the help of profiles that hold a complex networked multitude of data about humans (p. 113)." Park et al. (2015, p. 602) defines social surveillance as "the behavior of collecting and tracking others' information on social network sites", which actually misses the internalization of the user's feelings being surveilled. Tokunaga (2016) defines the term "interpersonal electronic surveillance" as the user's ability to provide their relationship behaviors through digital technologies. The different expressions in terms of the phenomena are nearly the same; however, the most stable one is from the part occurring within romantic interests, family, friends, classmates and colleagues (Trottier, 2012a). The availability of personal information on social network sites has opened the path to social surveillance.

2.2.2 Social Surveillance Consequences

The personal information material became a target by social surveillance practices. The nature of Facebook encourages disclosure, which also makes surveillance available at hand and makes privacy a difficult position (Trottier, 2012a). The users' personal information and all other visible activity on Facebook is the driving force of the production of gossip and rumors (Debatin et al., 2009). The audience (consuming-users) can reach the user profiles (producing-users), which contains disclosed information, at will, or the other way round. Surveillance is a vital construct for privacy concerns (Park et al., 2015). Making a person public is a habitual behavior in social surveillance when "users make part of their profiles and content visible to the public and to laterally observe what others are doing and posting (Fuchs & Trottier, 2015, p. 130)." Social network sites enable users to monitor others owing to the features of record-keeping and displaying it in social networks. Users are often uncomfortable by their constant observation, and the surveillance practices brings unwanted visibility, which leads to incurring nuisances (Trottier, 2012a). Individuals disclosing their personal information online normalizes surveillance practices and intensifies self-scrutiny, where people watch others and are being watched (Trottier, 2015).

With that connection, social privacy is a phenomenon consisted of the practices of information disclosure and social surveillance. The amount and kind of disclosed personal information to a specific audience is determined by the user. This decision-making carries awareness for empowering a self-attending individual to keep their own personal data under control. Young and Quan-Haase (2013) makes a definition of "social privacy" based on users' strategies to guard their personal information against privacy invasions. Social privacy awareness is a series of strategies and practices to manage and preserve personal information disclosed online from others' *surveillant* gazes. Being aware of social privacy affects users' more powerful standing in privacy violations occurring between individuals. Users should be aware of the importance of their personal information data online to preserve their privacy and prevent the misuse of disclosed information. Users with high levels of awareness then would have high concerns about social privacy. Therefore, social privacy awareness requires the knowledge of users to manage their personal information disclosed online (Litt, 2013), and also be conscious and informed about community-based threats and social network site policies (Dinev & Hart, 2005). Overall, awareness of social privacy is about the matter of being concerned, engaged, and knowledgeable about controlling and managing the aforementioned issues.

2.3 Visibility Strategies

Individuals who join social network sites reveal a great amount of personal information online, and they are not plainly forced to disclose anything. Conceiving users as passive participants on social network sites is the wrong approach to understand the practices of users in social network sites (Livingstone, 2008). Users often deliberately desire to maintain their disclosed personal information online. The growing concern about how their personal information is disseminated in the online environment is heavily expressed in the literature (e.g., Acquisti and Gross, 2006; Young and Quan-Haase, 2013). The practice of disclosing content on social media has impacted users to pursue some strategies to ensure their privacy (Debatin et al., 2009). University students, in particular, have developed strategies to mitigate their concerns about privacy and protect their personal information (Young & Quan-Haase, 2013).

Users' making personal information public increases users' visibility and how their information appears to others is considered to occur in such a manner that reflects their preferences and desires (Marwick, 2012). In other words, the content disclosed on social network platforms mainly needs a "desired audience" (Debatin et al., 2009) that the producing-user has willing to maintain control on what kind of personal information consuming-user is able to see. In either physical or online life, the personal information can be restricted to a certain audience consisting of a limited number of people who can view it. Facebook enables users to choose the audience of their content in terms of public, semi-public, or private settings. The public content aims for everyone to have access to the online public realm. Semi-public is desired for a specific audience such as friends, family, and so on. Private is the "only me" option.

2.3.1 Visibility and Privacy

The ever-growing features of social network sites have increased the level of privacy concerns (Livingstone, 2008). Although it can be only eavesdropping as in physical life, social media breaches can go as far as accessing personal information and finding ways of reaching it without intention (e.g., someone using a third common connection in their Friends network to monitor your profile). Therefore, the digital sphere has a great capacity for privacy intrusions on the ground that digital content is persistent, systematically-recorded and archived, replicable, searchable, and expressive of high visibility (Boyd & Marwick, 2011).

Brighenti (2007) attempted to create a general theoretical understanding of visibility for the social sciences. The state of "intervisibility" (p. 326), based on the reciprocal relationship of visuals, appears intrinsically in the natural settings of physical life (i.e., if a person sees you, you see them). As previously mentioned, that online audience is not limited to physical boundaries (i.e. walls, doors or any distance); this audience has become so hidden that individual users can no longer know the subject and time of the "gaze[7]." The majority of social network sites function in this way; for instance, on Facebook, users are not entitled to know whether they are being watched — and, if yes, by whom — due to "asymmetrical relations of visibility" (Brighenti, 2007, p. 338), and users cannot estimate when they are being watched because of the "asynchronicity of visibility", so that the act of seeing or being seen is no longer superior with one over another.

Visibility of personal information is involved in information disclosure practices, yet it is slightly different than information disclosure. Visibility is a more broad term also containing the practices of information disclosure. This thesis approaches visibility in the context of strategy. On that matter, visibility strategies are also a kind of similar phenomenon slightly different from disclosure. Visibility is the set of attempts of an individual to make itself visible. Visibility strategies are composed of user practices regarding the conditions of extent personal information becoming available for others to access.

Previous studies indicate that social network site users have a different variety of strategies to control personal information, like restraining the flow of disclosed information and managing privacy tools to set visibility (e.g., Boyd and Hargittai, 2010; Litt, 2013; Tufekci, 2008). Visibility of personal information online relates to social privacy awareness in terms of three main issues: First of all, users on Facebook can maintain the kinds and amount of personal information they disclose (e.g., Christofides, Muise, and Desmarais, 2009; Livingstone, 2008; Waters and Ackerman, 2011). Second, users can make the content visible to a specific limited audience performed through privacy tools (e.g., Acquisti and Gross, 2006; Boyd and Hargittai, 2010; Debatin et al., 2009; Litt, 2013). Third, the users can choose to give inaccurate or

[7] While "gaze" is more like a steady, fixed tool, "surveillant" (or "surveillance") implies a systematic mechanism of gaze.

incomplete information to achieve social privacy (Acquisti & Gross, 2006). Making combinations of strategical issues allows the possibility for some users to follow more than one strategy at the same time. Additionally, there is a fourth way of visibility strategy, which is not in the scope of this thesis[8], regarding the context of the disclosed information to be coded to a specific group of users — a subtle context in which, only a specific group of audiences can decode the "hidden meaning" that the people not intended to understand the message cannot comprehend the full implications of the message (Boyd & Marwick, 2011).

2.3.2 Personal Information Disclosure

Disclosure is an essential act for settling the public life of individuals that ensures social interaction in society in accordance with the level and kinds of information people reveal to each other (Joinson & Paine, 2007). The act of disclosing makes the previously unknown a shared piece of knowledge. Personal information disclosed by a user itself is self-disclosure, which is the disclosure of one's own self, or "making the self known to other persons (Jourard & Lasakow, 1958, p. 91)."

Individuals may proceed in producing and disclosing personal information on social network sites, despite the risks, because of perceived benefits. Livingstone (2008) mentions that young people are more likely to balance benefits and risks of privacy issues on social network sites. Young and Quan-Haase (2013) express the "privacy paradox", describing that users are willing to disclose personal information on social network sites in spite of a high degree of concern. The risks of disclosing personal information and privacy-related concerns on social network sites have become a centerpiece of discussion in the literature over the last years (e.g., Debatin et al., 2009; Lewis et al., 2008; O'Brien and Torres, 2012; Waters and Ackerman, 2011). It appears that the issue of disclosing personal information and controlling the visibility are surely a long privacy-related discussion, which attempts to examine the motivations of disclosed information, along with its risks and benefits.

The content disclosed on social media is created for different purposes with regards to the consideration of the reached audience. Users usually leave their personal information on social network sites to give ideas to others about who they are that is relatively similar to physical world that people have to imply who they are, what they do i.e. by their appearance. Whiting and Williams (2008, pp. 366, 367) describe several reasons why users would have a profile and disclose personal information on social network sites: communicatory utility to facilitate communication and share information with others; social interaction in order to interact and communicate with other people; information sharing to enjoy posting and sharing their personal information about themselves with others; and surveillance of others to watch and keep updated on what other people do.

[8] Hidden context may need to be examined in a qualitative study for a more sturdy analysis.

Debatin et al. (2009) remarks that the benefits of online social network sites outweigh the risks of disclosing personal information. Users pragmatically share their personal information because they expect benefits from public disclosure (Gross & Acquisti, 2005). In this respect, users are motivated to disclose owing to drawing attention to their social status, attention-seeking and making personal information about themselves visible to some audiences (Marwick, 2012). Being disclosed can be desirable based on the idea of "self-revelation" (Solove, 2002). Users may choose to expose their personal information in order to achieve publicity or, satisfaction or may want it because of validation from social ties (Trottier, 2012b). As users have different levels of concern about privacy issues, there is still the cost users' obligation to give up their privacy benefits (Joinson & Paine, 2007). Users have a trade-off understanding of disclosing personal information on social network sites, as the desire for disclosure rules out privacy or makes it less significant. Users' attitudes, practices, and strategies on online privacy asserts that many make the trade-off, knowing what they gain and what they lose (Boyd & Marwick, 2011).

Since social network sites have radically altered the means of the manner, style, and content of information and communication practices (Shao, 2009), Facebook functions as a multi-purpose platform for users to share personal life and communicate with friends, family, colleagues, acquaintances, and strangers within a network (Facebook, 2016). Facebook demands personal information from users, in the process of "creating your profile" when they sign up, such as their full name, birthday, current city, photographs, interests, and even home address, telephone number, and other kinds of important or less important details (Boyd & Hargittai, 2010). After creating a profile, users are able to connect with "Friends[9]" to interrelate with each other's personal information. Facebook users are able to connect with others by sending a Friend request. Facebook basically labels all contacts unequally as friends, and can cover a variety of people from celebrities to family members, or even friends not contacted in many years (Marwick, 2012). After the Friend request is accepted, users can correspond with each other and are allowed to access each other's profiles, which consist of numerous pieces of personal information shared online shared in asymmetrical and asynchronous ways (Brighenti, 2007).

Users can add personal information on Facebook in as much detail as possible, including their full name, age, date of birth, hometown, gender, contact details, multimedia uploads (photos, videos, sound, text), work, school, partner's name, relationship status, family members, religious views, political views, "liked" or "followed" pages, and joined groups (O'Brien & Torres, 2012; Tokunaga, 2011). Disclosing personal information on Facebook does not only occur on account profiles in the form of updating one's "status", but also in updates on

[9] Friends (with capital F) on Facebook refers to your connected network on Facebook including all the users you interact with, which is different than the everyday understanding of *friends*. The Facebook facility of Friends is always capitalized in the thesis to avoid any kind of misconceptions.

important events, picture sharing, and so on. The information can be also revealed in others' profiles, like in comments, on walls, on pages, in groups, etc. McKeon (2010) classifies the potential disclosed data on Facebook as self-explanatory information (e.g., name, gender, picture, birthday, contact info, wall posts, and photos); "Friends" in your network (the people with whom you are in connection); family members; city/location; place of birth; favourite books/music/movies; school or employment; religious views (or content related to it); and "liked"s pages, groups, persons, web pages, or any kind of entity possessing a "like button" on Facebook — and all the kinds of components helping extent the disclosure of personal information on Facebook.

The personal information disclosed online could leave some intellectual, emotional and relational traces behind, in which users are able to make inferences about other users (Cohen, 2008). Facebook, as a social network site, offers a high incidence of attainment to distributing personal information and to establishing interpersonal communication. The personal information disclosed online on Facebook is able to give a retrospective picture for monitored individuals, with all of its behaviors, expressions, and thoughts acquirable by viewing the news feed. Each item of personal information gives an impression to the consuming-user, although the personal information disclosed in the past may not accurately represent the present individual. Facebook facilitates an accumulation of that sort of information and enables users to create an identity, which can be traceable backwards. These sorts of facilitation and allowances serve to build constant meaning for the individuals.

As information disclosure is an indicator of concerns over awareness of social privacy, it is assumed that students who are more aware of social privacy will disclose less personal information compared to those who are not. Therefore, the following hypothesis is offered in terms of the amount of disclosed information of users on Facebook that is subject to change by the level of awareness:

H_1: *The amount of disclosed information of certain types will be positively associated with social privacy awareness.*

2.3.3 Technological Privacy Tools

Social network sites generally offer different privacy options functioning to adjust the users' level of visibility for different groups of audiences. Privacy settings are a special, mostly-used feature in social network sites for arranging the visibility of users' profiles, which can be controlled to maintain a public or semi-public profile[10] by courtesy of privacy settings (Boyd & Ellison, 2008). The personal information on Facebook could be disseminated to varying degrees

[10] Wyrwoll (2014, p. 15) describes different public levels that could be applied to most social network sites and user-generated media platforms: (i) *General public*, which has unlimited audience; (ii) *Unknown-limited public*, when audience is limited to unknown people e.g. "Friends of Friends on Facebook" (iii) *Known-limited public*, when audience is limited to people the user knows e.g. "Friends" on Facebook.

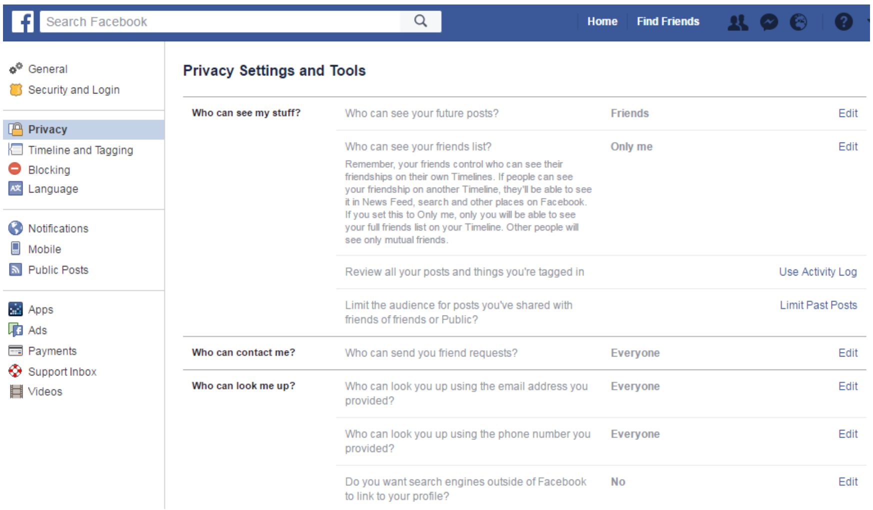
by users, who set them visible to a certain audience — such as those with whom the user has connected in their Friends network, some Friends within the network, or the whole public (Boyd & Hargittai, 2010). Debatin et al. (2009) emphasizes that restricting audience to a limited receiver through privacy settings is the most popular mechanism to control visibility of personal information. The user is empowered to reveal or mask the content from the public or Friends' network viewing by adjusting the setting of the audience's access. Privacy settings empower users to choose their audience in terms of who is able to access the different parts of their profile.

Nevertheless, privacy settings should not be reduced to one single function that only manages the size and range of the desired audience via levels of public access. Other kinds of instruments to maintain social privacy such as "removing tags" or editing contact the list are neglected in previous research (e.g., Boyd and Hargittai, 2010; Christofides et al., 2009; Lewis et al., 2008). These kinds of instruments take the concept of privacy settings a step further, and Litt (2013) hereby implies the term "technological privacy tools", describing all types of Facebook-platform-based-implications that are taken to maintain personal information privacy. The tools are in various forms such as "lists" allowing users to share information to their desired audience, "buttons" allowing users to delete shared contents, and other features like "untagging posts", images, and other kinds of information linked with users' identities. Similarly, Young and Quan-Haase (2013) describe strategies to maintain social privacy on Facebook. The applied strategies include that of excluding contact information, having a less detailed profile, "untagging" or removing names from photos, and rejecting or limiting Friend requests from strangers. Hence, the use of privacy settings is broadly conceived and should not be only reduced to audience limitation.

By the virtue of privacy tools on Facebook, users are able to maintain the visibility of their personal information in terms of what and how much to disclose and to whom to disclose. Facebook allows users to manage their privacy by controlling the level of visibility of their disclosed content[11] (see Figure 2.2). Changing visibility of personal information through privacy options on Facebook stipulates personal information not accessible by others who are outside the network and protects from surveillance of "unwanted" others, at least regarding the content. Acquisti and Gross (2006, p. 51) say that users are able to arrange the visibility of their profile, basically "who can read their profiles", as well as the searchability of their profile, or "who can find profiles through search features." Users remove or limit a viewing of an audience from the list, and Facebook builds a virtual boundary in between so the audience

[11] The privacy settings and tools by Facebook — 2017, June 7 retrieved from Facebook (2017): Who can see my stuff? (*Who can see your future posts? — Who can see your friends list? — Review all your posts and things you're tagged in. — Limit the audience for posts you've shared with friends of friends or Public.*) Who can contact me? (*Who can send you friend requests?*) Who can look me up? (*Who can look you up using the email address you provided? — Who can look you up using the phone number you provided? — Do you want other search engines outside of Facebook to link to your profile?*)

cannot access other's information any more (Litt, 2013).



Privacy Settings and Tools

Who can see my stuff?	Who can see your future posts?	Friends	Edit
	Who can see your friends list? Remember, your friends control who can see their friendships on their own Timelines. If people can see your friendship on another Timeline, they'll be able to see it in News Feed, search and other places on Facebook. If you set this to Only me, only you will be able to see your full friends list on your Timeline. Other people will see only mutual friends.	Only me	Edit
	Review all your posts and things you're tagged in		Use Activity Log
	Limit the audience for posts you've shared with friends of friends or Public?		Limit Past Posts
Who can contact me?	Who can send you friend requests?	Everyone	Edit
Who can look me up?	Who can look you up using the email address you provided?	Everyone	Edit
	Who can look you up using the phone number you provided?	Everyone	Edit
	Do you want search engines outside of Facebook to link to your profile?	No	Edit

Figure 2.1: Facebook's privacy settings, June 2017.

Boyd and Hargittai (2010) explain that a major privacy approach of Facebook is "network-centric", allowing users to control what personal information is shared with whom under the options — ranging from general to specific — of "Public" (anyone on or off Facebook), "Friends" (Friends network on Facebook), "Specific Friends" (only show to some Friends in network), and "Only me." Restricting disclosed information through privacy settings on Facebook can either increase or decrease visibility and exposure of personal information disclosed online (Trottier, 2012b; Tufekci, 2008). For instance, when a user does not want to show a photo to certain users (i.e., a user among friends, or just the public); they can use these privacy tools to restrict access to that photo (Litt, 2013). According to Facebook (2017), there are multiple ways for users to manage their privacy via the the privacy tools offered by Facebook. The "audience selector" tool (see Figure 2.2) functions to select who users share their status, photos, and other kinds of content they post. This tool is available in multiple places, such as privacy settings page and privacy shortcuts. The tool updates itself in multiple places when users change in audience selector tool in one place. After users share a post, they have option to change the audience which the content is shared with. When a content is posted to another user's Timeline, the other user can control the audience. In addition to that, the users tagged in the post might see the content along with their Friends.

Lewis et al. (2008) conclude that students are more likely to have a "less public profile" on social network sites if they are relatively more active than others. Litt (2013) asserts that active users on Facebook are more concerned about their privacy due to a greater disclosure of information. Boyd and Hargittai (2010) specify that users who use Facebook more frequently

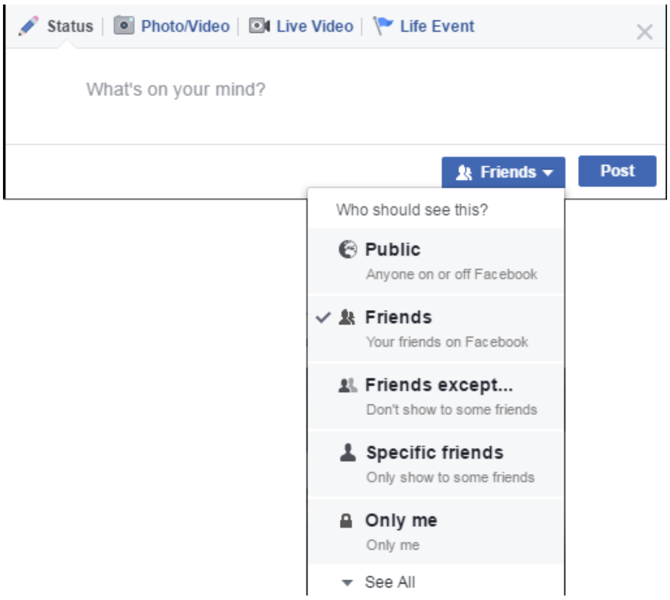


Figure 2.2: Audience selector tool in Facebook, June 2017.

or intensively have the inclination to use technological privacy tools more frequently. Furthermore, Boyd and Hargittai (2010) remarks that privacy settings are favorable to those who regularly use Facebook. There is a correlation between frequency of Facebook use — being active (i.e. posting content on Facebook) — and adjusting privacy settings (managing who can see the information). The measurement of "Facebook intensity" is based on users' levels of engagement involving all kind of actions to use Facebook in an effective way (Ellison, Steinfield, & Lampe, 2007), including these: how long user has been on Facebook, how many hours user spends checking its news-feed per day, how often user goes on Facebook, how often user shares content on Facebook, how many Friends user has, and what user's connections with its Friends are in terms of degrees (i.e., having a high or low degree to a group of Friends). As use of privacy tools is a strong indicator of concerns over social privacy, it can be assumed that students who are active on Facebook will increase the use of privacy tools compared to those who are not active, and they are also more aware about social privacy. In this vein, the following hypothesis is offered in terms of users' practices of technological privacy tools being subject to change by degree of activeness on Facebook:

$\mathbf{H_{2a}}$: *Intensity of Facebook use will be positively associated with social privacy awareness.*

$\mathbf{H_{2b}}$: *Intensity of Facebook use will be positively associated with the practice of technological privacy tools.*

O'Brien and Torres (2012) state that more than one-third of users have used the privacy related tools on Facebook to have higher control over their accounts. Therefore, users who have a high level of social privacy awareness are more likely to control through all settings, assuming that the users who have changed visibility via privacy tools are more aware of social privacy than those who have not used privacy tools (Acquisti & Gross, 2006; Litt, 2013; Tokunaga, 2011). Since higher levels of public restriction are positively associated with awareness of social privacy, the following hypothesis is offered:

H$_3$: *The use of technological privacy tools employed in order to diminish visibility will be positively associated with awareness of social privacy.*

Facebook has a considerable number of users who have an open profile sharing their personal information with "strangers" (Tuunainen et al., 2009). Although most Facebook users claim that they know how to control the visibility of their profiles (Acquisti & Gross, 2006), many users do not even change default recommended settings on Facebook, which are offered at registration. The default recommended privacy setting of Facebook is set to Public, which makes disclosed content open to all users and possibly even non-users in the network. Default privacy settings of Facebook do not make users' personal information more private; instead, it makes the user's name, profile picture, current city, gender, "liked" pages and groups, connected Friends, etc. publicly visible (Waters & Ackerman, 2011). On that issue, the following hypothesis is offered in terms of users who have changed default recommended privacy settings and are more aware about social privacy:

H$_4$: *Having changed their default recommended privacy settings on Facebook will be positively associated with awareness of social privacy.*

2.3.4 Completeness and Accuracy of Information

As disclosure is an essential act for public life of individuals (Joinson & Paine, 2007), yet "decreasing profile visibility through restricting access to Friends" (Debatin et al., 2009, p. 103) — privacy settings that give control to users to limit the visibility of their personal disclosed information, helping users avoid social surveillance and achieve social privacy — is one of the most common strategies to avoid an unwanted audience. It is stated that full prevention against privacy violations is not possible. Some kinds of privacy violations occur not because of the lack of privacy tool uses, but because the privacy that is managed via privacy tools can be still breached through direct networks on Facebook (Waters & Ackerman, 2011). For instance, user profiles can be monitored by another user (Friend of Friend or gazing from a Friend's profile) who is not a Friend of that user.

The visibility of disclosed information might not be suffice maintained only by the use of privacy tools; sometimes, users can give false or inaccurate information, thinking that they can maximize their privacy (Trottier, 2010). Das and Kramer (2013) identifies the situation as self-censorship practices on Facebook, which is "the act of preventing oneself from speaking (p. 120)." Decreasing the volume of disclosed information or altering the accuracy of information is a set of practices maximizing the privacy of content. Changing the accuracy of information or not giving complete information could prevent the "[risk of] unknown or potentially inappropriate audiences gain access to a user's content (Das & Kramer, 2013, p. 121)."

The fact that Facebook users are content (Debatin et al., 2009) in managing unwanted audiences by limiting their overall profile and specific areas of content visibility through technological privacy settings Tufekci (2008) conveys that users tend to believe that Facebook privacy settings are convenient and effective enough for keeping unwanted audiences away. However, this may lead the "illusion of control", which motivates users to share more and makes them unaware of who else could have access to that personal information, such as "eavesdroppers" and "gossipmongers" (Boyd & Marwick, 2011). On that matter, some users might be concerned that their content could be leaking to other profiles of whom the user does not desire. So changing the meaning of content is a frequently used strategy for the sake of social privacy. Making the content inaccurate can be a valid strategy as it leads to a variety of interpretations by other users on social network sites. For instance, a message encoded and posted on Facebook can only be decoded and interpreted by the related audience of Friends and can help decrease the level of vulnerability of social surveillance. The decoding of message can prevent the social surveillance happened through profile breaches from a third person. Eventually, students could alter or limit the "authenticity" of their personal information by giving inaccurate or incomplete information in order to increase the level of control of their personal information, instead of just relying on privacy tools to adjust the user's visibility. This could help users express themselves without the risk of intrusion by unwanted audiences.

2.4 Negative Social Network Site Experiences

Thanks to the convoluted lives of both digital and physical spaces (Cohen, 2008; Trottier, 2012b), activities conducted online could lead to negative consequences in physical life. Young and Quan-Haase (2013) remark that negative social consequences have a prominent impact on users' behaviors concerning privacy practices. For instance, young people often have trouble with their privacy on their social network site accounts. They sometimes face problems of losing control of the account (e.g., having an undesirable post shared by other users or being hacked). In addition, disclosed personal information of users might have unknown negative effects in the future, or the information is discovered for an evidence against users (Nissenbaum, 2009). Individuals do not want "unintended consequences" (e.g., Debatin et al., 2009; Livingstone, 2008; Waters and Ackerman, 2011) or their personal information to affect their lives unfavorably and affect their right to privacy.

The practical implementations of social surveillance can be various, and some are derived from experiences such as stalking, harassment, watching, spamming, creeping, gazing and looking (Govani & Pashley, 2005; Marwick, 2012). These practical implementations can end up with some consequences such as embarrassment or reputation damage, harassment or stalking, and naming and shaming practices (Boyd & Ellison, 2008; Debatin et al., 2009; O'Brien & Torres, 2012; Trottier, 2012a). Moreover, Govani and Pashley (2005) describes that users may be abused or faced with stolen identity because the information given to Facebook is personally identifiable. Similarly, the consequence of personal information being disclosed online could be seen by an audience which the information was not intended for. The personal information disclosed online could be used in the future by of state government and future employers to judge a user's character.

Debatin et al. (2009) and O'Brien and Torres (2012) found that the large majority of students are content to disclose personal information although they are aware of possible consequences of privacy violations that are caused by sharing personally identifiable information in an unbounded network. Tuunainen et al. (2009, p. 5) describe "loss of privacy and control over personal information may cause damages that are socially irreparable: losing face among friends, revealing secret information, making social blunders, or simply giving a wrong impression"; privacy concerns on social network sites may cause serious social problems that have significant losses for the individual due to the issue of unauthorized access. Furthermore, they imply that these damages can easily become more serious when the audience includes people with whom the user has to casually interact in the physical world.

Social surveillance may incur fundamental consequences for the user in physical life because of the pervasive condition of social media surveillance normalizing surveillance practices (Trottier, 2012a). Since the ways to communicate between people are flowing either in physical or online space, all the actions taken in physical space can have significant consequences online, or vice versa (Cohen, 2008; Trottier, 2012b). The previous experiences on social privacy violations emerge from social surveillance practices where individuals monitor others in the interest of gathering personal information. The reasons for more consequences are that social media is a more convergent place, in which the accessibility and searchability of personal information makes privacy boundaries more vulnerable.

Users are likely to change their privacy settings if they had an experience, that of a privacy violation or a negative social consequence caused by information disclosure on Facebook, at first hand or heard second-hand (Debatin et al., 2009; Young & Quan-Haase, 2013). Litt (2013) specifies that users are motivated to raise their privacy settings strategies and to reconsider their disclosure patterns when they experience "privacy turbulence", which implies a privacy violation to one's desired boundaries.

H_5: *Users who have previously heard negative situation will be positively associated with social privacy awareness.*

H_6: *Users who have previously experienced negative situation will be positively associated with social privacy awareness.*

3 Method

3.1 Survey Design

Surveys are quantitative instrument used in social science and are designed to ask close-ended questions, which usually follow a question/answer design, to a large number of people (respondents) regarding their opinions, behaviors, thoughts, and so on (Neuman, 2014). A quantitative cross-sectional survey research method is chosen in this thesis to be able to explore general patterns of social privacy awareness on Facebook. Surveys are convenient method to gather large amounts of data representing the general characteristics of a large population in a short time period. Internet surveys are flexible and successful for collecting high amounts of data from a sample population in a short time and at a low cost (De Leeuw, 2008).

De Leeuw (2008) explicates that a survey questionnaire conducted on the Internet can have complex questions, but they should be entirely self-explanatory. Internet surveys can have the advantage of using visual aids. Moreover, they can be less intrusive and more private than other types of surveys (e.g., face-to-face, telephone).

3.2 Population and Sampling

Simple random sampling, the simplest form of probability sampling, is a selection process that enables the choice of possible subsets of size n from a population of size N, where every person in the population has the same probability ($=n/N$) (Lohr, 2008). Since observations of a sample of the population are smaller than the whole population, and it is hard to establish the representativeness of the sample, which is known as the sampling error, the probability sampling helps diminish coverage error and provide more reliability (Lohr, 2008). In order to avoid bias towards any other particular background, any method leading to selective sampling will be refrained. Selective sampling (i.e., approaching respondents on the street or any kinds of snowball sampling) is avoided to remove bias from any social background and the researcher's own personal background. Since online surveys do not provide easy randomization (Riffe, Lacy, & Fico, 2014), the online survey link is offered in various social network sites and different university websites, so as not to constrain sampling to a specific environment.

The sample population should be representative of the study population (Roberts, 2006). The representative sample in the survey is university students in the Netherlands who have

Facebook profiles. According to CBS (2012)[1], the total students attending full-time, part-time, and exchange based university education is 669.041[2]. According to Pew Research Center (2017), Facebook is the biggest and most famous social network site in the world with 1.86 billion monthly active users. Facebook is the most visited and engaged social media platform among ages 18 to 34 (Clark, 2016). The defined place — the Netherlands — is chosen to make this thesis more feasible within a defined time frame.

University students in the Netherlands are employed to define which respondents are allowed to complete the survey/conduct survey. A respondent student studying in either WO or HBO universities is sufficient to answer the survey questionnaire. Having a Facebook profile is a prerequisite as well, which is asked in the beginning of the survey questionnaire. Those who do not have Facebook profiles or have not been on Facebook before cannot proceed to answer beyond the demographic data questions. The amount of people who stated that they have a Facebook profile was 96.16%, and only 2.84% (five respondents) stated not to have a Facebook profile, which lead them to the second prerequisite question: if respondents have ever had a Facebook profile. To this question, three respondents said “Yes” and proceeded to the rest of the survey questions, and 2 respondents said "No", and their responses were recorded and the survey ended.

De Leeuw (2008) depicts that nonresponse rate of a survey is based on the failure to measure sampled units. Nonresponse, where units are selected into the sample but not measured, is different than coverage error, where units do not have the chance of being selected in the sample (e.g., for that case, the students who do not have Internet connection — it is unrealistic in the Netherlands but it may exist). A nonresponse rate cannot be calculated in Internet surveys because the number of people who had the opportunity to participate in the survey is unknown.

3.3 Data Collection and Analysis

This thesis draws cross-sectional survey data, which collects the data at only a point of time, collected from different respondents in the sample population. The university students were invited through social network sites (Facebook, Twitter, and LinkedIn) groups, pages, and other kinds of online platforms. The invitations were distributed equally (e.g., not too many survey links on just one social network site, like Facebook, group was shared), for the purpose of avoiding bias (e.g., the respondents might be aggregated from one university).

[1] Statistics Netherlands known as *Centraal Bureau voor de Statistiek* (Central Bureau for Statistics). http://www.cbs.nl/

[2] Higher education in the Netherlands has two types of institutions: WO (*wetenschappelijk onderwijs*) literally meaning “scientific education” and HBO (*hoger beroepsonderwijs*) meaning “applied sciences”. According to CBS (2016), 245.322 in WO and 423.719 in HBO students with sum of 669.041.

The survey is held online via Qualtrics[3], which has an institutional participation with Erasmus University Rotterdam. The survey starts with clear consent from the participant. It was ensured that the data acquired from the survey would stay anonymous and only be used for academic purposes. The anonymity of respondents, who participated in the online survey via Qualtrics, is secured. On average, it takes 5 to 6 minutes to complete the survey. The survey had been dispersed online through social network site accounts (Facebook, Twitter, and LinkedIn). The questionnaire was available for 30 days.

The survey link, inviting qualified people to participate in the research, included a short description about the thesis research, defining the most-used privacy terms and ensuring the anonymity of individual responses. The survey was conducted in May 2017 with a total of 176 students, and only 169 respondents finished the survey ($N = 169$) from different backgrounds with the response rate[4] of 96%. The information about the students who refused to participate is unknown. The bias against students who are not frequently online or not visiting such groups and pages might be existing because the survey is administered on the Web. The total number of respondents (N) is suitable for the statistical analysis with a 95% confidence interval, which is based on both the sample size and the variance of the measurement.

The survey results are statistically analyzed by R language[5] on the OS X operating system. The visualizations of the descriptive and inferential results derived from the survey data are created by R and R's `ggplot2` package[6]. The study variables are comprised by continuous (i.e. *age*), and categorical (i.e. *yes/no*). The survey data is extracted from the Qualtrics as CSV format[7].

In the survey data, Cronbach's alpha (α) was used for reliability, and PCA Factor analysis was used for validity analysis. Multiple t-test, ANOVA, and further post-hoc tests are implemented. For the constructs extracted from the reliability and validity analyses, several multiple regression models were presented along with their assumptions checks.

3.4 Operationalization

The operationalization deduces conceptualized variables into measurable units for the questionnaire in the survey (Neuman, 2014), and answers the hypothesis derived from different theoretical approaches. The survey questionnaire contains 18 primary questions, and 2

[3] Qualtrics is an online survey tool which people can easily use to create and disperse their own surveys on the Internet: http://www.qualtrics.com/

[4] "Response rate" is the percentage of how many people completed the survey calculated as: (Total Finished of Responses / Total Responses) × 100

[5] An open source programming language mainly used for statistics and also other multiple purposes such as network analysis, data analysis, and machine learning: http://www.r-project.org

[6] For the documentation, see https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf

[7] Comma Separated Values, or CSV, is a plain text format document storing data in a structured tabular form, where the each record is delimited by commas.

prerequisite questions. The questions classes are demographic details, intensity of Facebook use, awareness of social privacy, personal information disclosure, completeness and accuracy of information, technological privacy tools, and previously negative experiences. Categorical (nominal, dichotomous, ordinal) and continuous (interval) variables are all used. The survey begins with demographic questions asking gender, age, nationality, university department, and level of education. Gender is asked in three options: male, female, and other with the objective of gender neutrality. The respondents' nationalities is asked instead of asking their race or ethnicity. The other demographic questions inquire about students' university department and education levels. Following that, whether respondents have a Facebook profile or not appeared as a prerequisite question. If the respondent says "Yes", the survey continues as usual. If the respondent says "No", a second question is asked whether the respondent has had a Facebook profile. If "Yes", survey continues as usual; if "No", the survey is finished for that respondent, whose demographic details are only recorded.

Intensity of Facebook Use. This part of the questions solicits the amount of time spent on Facebook (in other words, the activeness of respondents on Facebook). First of all, it asks how many years respondents have been on Facebook. The answer can be chosen in the range between 2004 to 2017. The following questions are followed by Horváth, Bogaerts, Sijtsema, and Demeyer (2014); firstly, examining the number of hours of respondents checking news feed on a daily basis, which is ranged from 0 to 12 or more hours, appeared with the answers of 0 to 1, 1 to 3, 4 to 8, 9 to 12, and more than 12. Second, the frequency of going on Facebook is asked with 6-point rating scale answer items, which are descendingly sorted as "more than twice a day", "once a day", twice a week or more", "once a week", "once a month", and "never." Third, the question asking, "On average, how often do you share content on Facebook?" (like updating status, adding photo, check-in somewhere, etc.) has the same answers as the question before. In addition, the following item questions are adapted from Young and Quan-Haase (2013), which especially focus on the Friends feature of Facebook. The first question asks respondents the number of connected Friends on Facebook (approximately). Second, a matrix-based degree table is formed to ask how many connected Friends of respondents are considered as "Close friends", "Acquaintances", "Distant friends", and "People only met on Facebook" within the degrees of "Low", "Moderate", "High", and "None".

Personal Information Disclosure. This is a matrix-type of question asking, "What kind of personal information do you have shared in your Facebook profile, and how complete and accurate is it?" with the nominal answers of "I share this information completely and accurately", "I share this information, but it is not complete or accurate" and "I don't share this information." The following question is inspired by multiple researchers (Acquisti and Gross, 2006; Tufekci, 2008; Tuunainen et al., 2009; Christofides et al., 2009) in order to understand what kind of information users disclose and on what basis. The disclosed personal information items are various (e.g., political views). This question is the combination of two variables: first, measuring the amount of personal information on the basis of whether it is shared or not, and second, completeness and accuracy of information. The measurement of completeness and accuracy of disclosed personal information is particularly applied by Acquisti and Gross (2006) to measure the accuracy of personal information, which users disclose online. As previously mentioned, a question asked what kind of personal information users provide on Facebook. Acquisti and Gross (2006) ask whether the information they disclose is accurate or not: "How accurate is the information you provide?" Basically, this is intended to know whether users alter the accuracy of their information as a strategy of preserving social privacy or not.

Technological Privacy Tools. The issue of use of technological privacy tools is introduced in two parts. The first part, the matrix table question "How often do you perform the following things on your Facebook account?", is largely applied from Litt (2013) and Young and Quan-Haase (2013), and is actually measuring privacy protection strategies which a number of strategies respondents employ to protect their personal information data. The main question asks how often users perform the things related with privacy tools on Facebook because asking frequencies is a convenient question form rather than directly asking what their privacy protection strategies are, which would be too obvious and could have led respondents' thoughts towards the issue. The answer measures are applied from Young and Quan-Haase (2013) with the items asked in a 5-point Likert scale (from 1 = *Never* to 5 = *Always*) along with some questions (i.e. "I have untagged myself from images and/or videos posted by my contacts"). Some question items are adopted by Wisniewski, Knijnenburg, and Richter (2017), i.e. "I have gone 'offline' on Facebook chat." In this question, and also the others, the 5-point Likert scale is the most preferred instead of other alternatives, like 7-point. A larger point Likert scale may cause an increment in "frustration level" (Buttle, 1996) because more answers cause users to take more time to make a decision, even though high points seems that there is more variance in terms of measurement. The question items look diversified, and are thus adopted and applied from different resources. These 11 different items in the question could be distributed into five different categories as Litt (2013) describes: (1) change privacy settings; (2) delete people from network/Friends lists; (3) untag photos; (4) limit certain updates to certain people; and (5) delete others' comments from their profile.

A second set of questions about technological privacy tools are individual dichotomous "Yes" or "No" type of questions. Since the default recommended privacy settings of Facebook are set to Public, which makes user profiles and their content open to all users in the network, even strangers, it asks for a "Yes" or "No" response to the statement, "I have changed the default privacy settings recommended by Facebook." In addition to that, the question "Are you aware that if you have joined some network and you haven't changed your privacy setting, all members from the same network can see your profile?", which is applied from Tuunainen et al. (2009), intending to measure whether users do not use privacy settings out of ignorance or lack of interest (Govani & Pashley, 2005) with the same, “Yes” or “No”, answers like in the previous question.

Awareness of Social Privacy. The question “To what extent do you agree with the following on your Facebook account?” is implemented from the work of Child and Starcher (2016); the question actually measures social (interpersonal) surveillance in which items and word choices are originally guided from the qualitative work of Trottier (2012a). Some items in the question are complementary, testing to what extent the respondents’ answers are reliable and honest, e.g. “I do not give much thought to whether people are actively monitoring what I post” and “I do not think about who may be constantly monitoring my Facebook page” are quite the same. The items are answered in a 5-point Likert scale ranging from 1 = *Strongly disagree* to 5 = *Strongly agree.* The item asked in the question “How often do you perform the following things on your Facebook account?”, which is situated in the variable of Technological Privacy Tools, is adapted from Tuunainen et al. (2009): “I worry that I will be embarrassed by wrong information others post about me on Facebook” asked in a 5-point Likert scale like the previous question (from 1 = *Never* to 5 = *Always*). Furthermore, social privacy awareness is a broad concept; as it is described in the theory chapter, the other variables could imply it. For instance, Technological Privacy Tools and Personal Information Disclosure variables involve the ability to control personal information in one’s profile so that the other variables could have an influence on the results.

Negative Social Media Experiences. These questions inquire about previous experiences of privacy violations. The questions are inspired from Litt (2013). The first question is “Have you ever heard any bad situation happen because of sharing personal information online?”, regarding whether users have witnessed any negative situations due to the disclosure of content online, and the second question is “Have you ever experienced any bad situation happen because of sharing personal information online?”, asking whether users had personally experienced a negative situation due to disclosing content online.

4 Results

4.1 Descriptive Statistics

4.1.1 Demographics

A total number of 169 student respondents finished the survey out of 176 students for this thesis study with a 96% response rate. The demographics of the students who finished the survey is presented in Table 4.1. Although the survey is aimed to represent all genders, more females (scored an average of 58.52%) than males (scored an average of 40.34%) participated in the survey (and Other reached 1.14% which is removed from further analysis due to lack of representation in the sample). Most of the respondents are in the age group of 19 to 25 (81.14%). Following that, students were asked their level of education including the following options: (1) 1$^{\text{st}}$ year (17.61%); (2) 2$^{\text{nd}}$ year (17.05%); (3) 3$^{\text{rd}}$ year (22.16%); (4) Masters (35.80%); (5) Other (7.39%). Most responses are recorded as Masters students (35.80%) and 3$^{\text{rd}}$ year students (22.16%), who cover more than half of the respondents.

	N or *M*	%	*SD*
Gender:	169		
Males	71	40.34	
Females	103	58.52	
Other	2	1.14	
Age	22.61		3.42
16—18	8	4.57	
19—25	142	81.14	
26+	25	14.29	
Level of Education:			1.23
1$^{\text{st}}$ year	31	17.61	
2$^{\text{nd}}$ year	30	17.05	
3$^{\text{rd}}$ year	39	22.16	
Masters	63	35.80	
Other	13	7.39	

Table 4.1: Sample demographics

Figure 4.1 shows the waffle chart of respondents' nationalities, which are the eight most common nationalities of the respondents who participated in the survey. The most common nationalities are the Netherlands (46.83%), which covers almost half of the respondents, followed by Turkey (12.70%), Germany (8.73%), China (7.94%), Italy (7.94%), Greece (5.56%), Poland (5.56%), and France (4.76%). Next, Figure 4.2 displays a waffle chart of respondents' academic disciplines. Despite that the survey is aimed to be dispersed to all students coming from different department backgrounds, it is found that the majority of the students (85.23%) study in a social sciences-related department and only a minority of students (13.64%) study in natural sciences-related department[1]. As the question aiming to know students' academic disciplines was collected as open-ended consisting of character and numeric data, the social and natural science studies are manually recoded by the researcher[2]. Overall, for these two visualizations, the waffle charts were preferred over common pie charts because estimating the proportion of respondent backgrounds could be messy to read in a pie chart, especially when estimating a single proportion or comparing a small number of proportions (Spence, 2005).

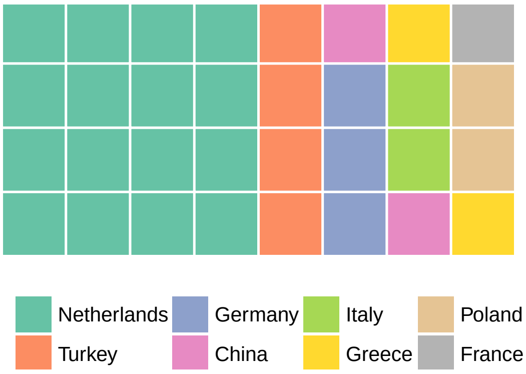


Figure 4.1: Waffle chart nationality

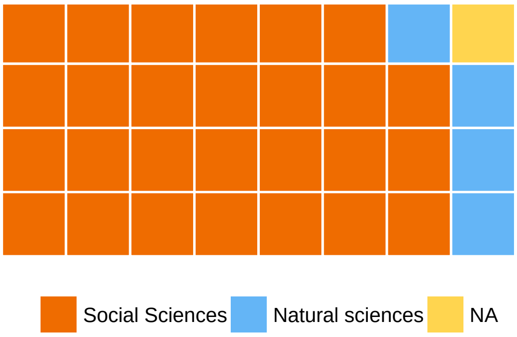


Figure 4.2: Waffle chart academic discipline

4.1.2 Dependent variables

Intensity of Facebook Use

Table 4.2 shows the frequency of respondents' intensity of Facebook use with the sample demographics. The year of being on Facebook shows the frequency since the users registered on the platform. The majority of the respondents have joined Facebook in the years of 2008 (20.22%), 2009 (19.1%), and 2010 (15.00%). The average of hours per day of checking the news feed on Facebook points out that the majority of respondents reported that they are checking 1-3 (50.39%) and 0-1 (38.25%). Next, the number frequency of respondents going on Facebook specifies that the majority of respondents reported that they go on more than twice a day (81.35%), while the "never" response stays at zero. The respondents' amount of shared content (i.e., update status, add photo on Facebook) per day average implies that the majority of respondents reported that they are checking 1-3 (50.39%) and 0-1 (38.25%). Surprisingly, the majority of respondents reported that they have more than 500 Friends (43.38%), yet the items seem to be dispersed as gradually lowering. Lastly, the level of connection with Friends

[1] Missing values (NA) is 1.14%.

[2] A full list of study departments can be found in Appendix A.

illustrates respondents' levels of connection with their Friends on Facebook by the term of degrees. It was reported that 43.27% of respondents think that they mostly have connected with Acquaintances at a Moderate level; 44.44% respondents think that they mostly have connected with Close friends at a Low level; and only 21.87% reported that they have no Distant friends. The last group, the People Only Met on Facebook, seems the most interesting; the respondents in this group reported no moderate degrees, as 41.52% respondents said "High", when 46.78% respondents said "None."

	M	*SD*
Intensity of Facebook Use		
How long have you been on Facebook? Since... (2004 — 2017)	6.88	2.46
How many hours do you spend on checking your news feed per day?	1.76	0.73
How often do you go on Facebook?	5.65	0.77
How often do you share content?	2.25	0.97
How many Friends do you have?	3.83	0.81
Level of connection with your Friends:		
Close friends	3.23	1.28
Acquaintances	3.16	0.72
Distant friends	3.04	0.82
People only met on Facebook	2.24	0.56

Table 4.2: Summary statistics for Intensity of Facebook Use

Personal Information Disclosure

The set of personal information disclosure questions were aimed to show what kind of information Facebook users share in their profiles or statuses. The abbreviations were used for aesthetic reasons: **NS** means "I don't share this information", **S-BNC/I** means "I share this information but it is not complete or accurate", and **S** means "I share this information completely and accurately." Table 4.3 shows the amount and kinds of personal information disclosed on Facebook by gender of the respondents. The most remarkable items were reported as follows: Full name (M = 1.24, SD = 0.50) is generally shared in both male (83.33%) and female (76.70%). Date of birth (M = 1.40, SD = 0.71) is common to share completely and accurately like full name; there is a high sharing rate for both male and female, 77.27% and 71.84% respectively. 63.64% males did not share their e-mail address (M = 2.52, SD = 0.80), while females were at 76.70%. 15.15% males reported that they share telephone numbers (M = 2.82, SD = 0.53) completely and accurately while only 1.94% females do this. Home address (M = 2.94, SD = 0.31) shows the highest proportion of not shared information among the other items for both male (1.52%) and female (1.94%). Political views (M = 2.64, SD = 0.67) are not a very shared topic, only 13.64% male and 9.71% female shared this information completely and accurately. Religion (M = 2.74, SD = 0.62) is another not very shared topic;

78.79% of males and 86.41% of females did not share it. Photos of you ($M = 1.37$, $SD = 0.56$) was mostly shared personal information for both males (66.67%) and females (67.96%), and only 27.27% males and 29.13% females reported that they share this information but it is not complete or accurate. Opinions about job, school, and family ($M = 2.65$, $SD = 0.69$) were reported, and 72.73% of males and 82.52% of females did not share this information.

There was a general trend that respondents chose not to share information as complete and accurate except a few, such as full name. Additionally, the sensitive information, such as telephone number and home address did not shared completely and accurately. Females tend not to share personal information more than men in all areas.

	NS		S-BNC/I		S	
	Male	Female	Male	Female	Male	Female
Full name	2 (3.03%)	4 (3.88%)	9 (13.64%)	20 (19.42%)	55 (83.33%)	79 (76.70%)
Date of birth	7 (10.61%)	15 (14.56%)	8 (12.12%)	14 (13.59%)	51 (77.27%)	74 (71.84%)
Hometown or City	8 (12.12%)	20 (19.42%)	11 (16.67%)	8 (7.77%)	47 (71.21%)	75 (72.82%)
E-mail address	42 (63.64%)	79 (76.70%)	8 (12.12%)	6 (5.83%)	16 (24.24%)	18 (17.48%)
Telephone number	53 (80.30%)	98 (95.15%)	3 (4.55%)	3 (19.42%)	10 (15.15%)	2 (1.94%)
Home address	61 (92.42%)	100 (97.09%)	4 (6.06%)	1 (0.97%)	1 (1.52%)	2 (1.94%)
Relationship status	45 (68.18%)	63 (61.17%)	8 (12.12%)	14 (13.59%)	13 (19.70%)	26 (25.24%)
Biography	43 (65.15%)	79 (76.70%)	13 (19.70%)	14 (13.59%)	10 (15.15%)	10 (9.71%)
Family members	36 (54.55%)	44 (42.72%)	19 (28.79%)	39 (37.86%)	11 (16.67%)	20 (19.42%)
School or employment	10 (15.15%)	7 (6.80%)	15 (22.73%)	29 (28.16%)	41 (62.12%)	67 (65.05%)
Political views	45 (68.18%)	82 (79.61%)	12 (18.18%)	11 (10.68%)	9 (13.64%)	10 (9.71%)
Religion (or related to it)	52 (78.79%)	89 (86.41%)	5 (7.58%)	7 (6.80%)	9 (13.64%)	7 (6.80%)
Sexual orientation	42 (63.64%)	82 (79.61%)	5 (7.58%)	5 (4.85%)	19 (28.79%)	16 (15.53%)
Partner's name	49 (74.24%)	75 (72.82%)	5 (7.58%)	7 (6.80%)	12 (18.18%)	21 (20.39%)
Family's name	27 (40.91%)	57 (55.34%)	10 (15.15%)	13 (12.62%)	29 (43.93%)	33 (32.04%)
Photos of you	4 (6.06%)	3 (2.91%)	18 (27.27%)	30 (29.13%)	44 (66.67%)	70 (67.96%)
Photos of your friends	14 (21.21%)	16 (15.53%)	18 (27.27%)	27 (26.21%)	34 (51.52%)	60 (58.25%)
Your travelling status	27 (40.91%)	55 (53.40%)	14 (21.21%)	20 (19.42%)	25 (37.88%)	28 (27.18%)
Opinions about job, school, family	48 (72.73%)	85 (82.52%)	7 (10.61%)	8 (7.77%)	11 (16.67%)	10 (9.71%)
Places you visit	32 (48.48%)	50 (48.54%)	14 (21.21%)	26 (25.24%)	20 (30.30%)	27 (26.21%)
Favorite music, book, movie etc.	24 (36.36%)	51 (49.51%)	15 (22.73%)	28 (27.18%)	27 (40.91%)	23.30%)
Important Life Events	27 (40.91%)	51 (49.51%)	20 (30.30%)	25 (24.27%)	19 (28.79%)	27 (26.21%)

Table 4.3: The amount and kinds of personal information disclosed by gender.

Next, Figure 4.3 shows the stacked bar graph of frequencies of status of each of the twenty-two items. The most remarkable items were as follows: Full name (77.27%), Date of birth (71.59%), and Hometown or City (69.32%) show high levels of sharing. School or employment (62.50%), and Photos of you (65.34%) show moderate levels of sharing this information completely and accurately. E-mail address (69.89%), Telephone number (86.93%), Home address (92.61%), Political views (73.30%), and Religion (81.25%) show high levels of not sharing this information completely and accurately.

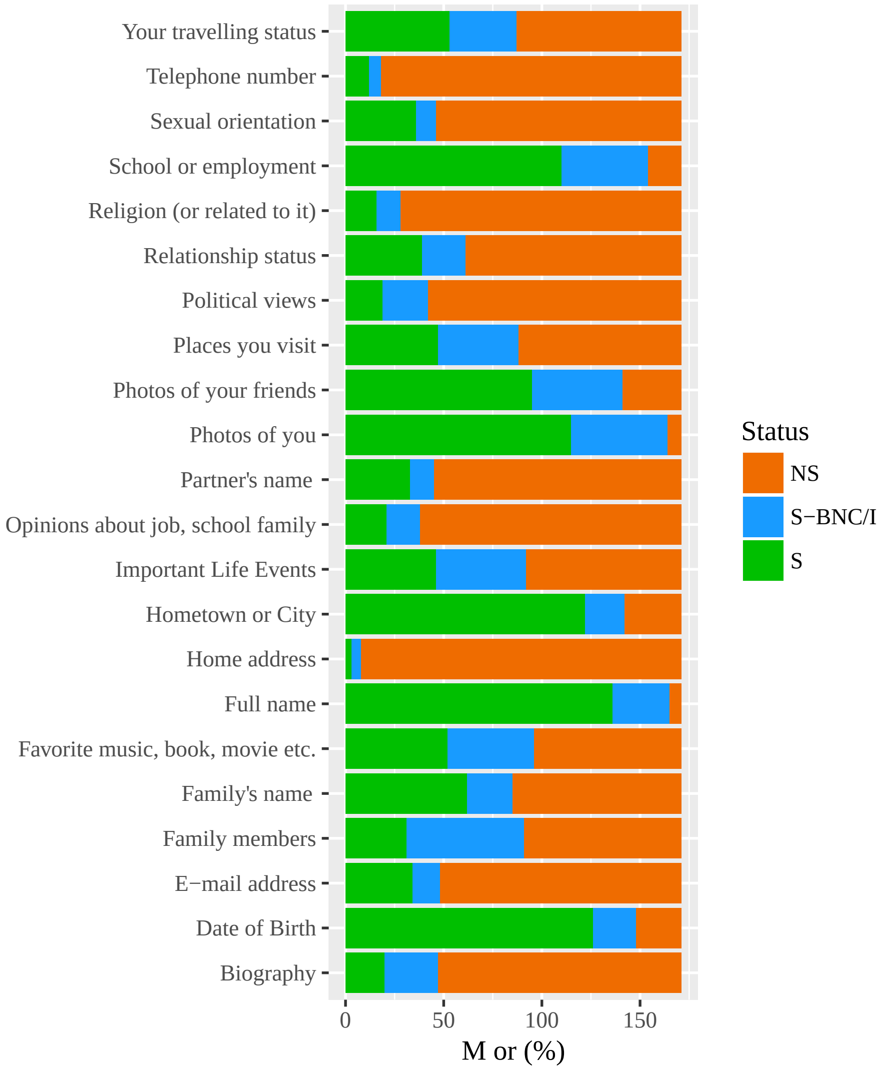


Figure 4.3: Bar Graph for Kinds of Information Disclosure on Facebook

Technological Privacy Tools

The set of technological privacy tools questions were aimed to show what each kind of privacy tools or settings share and to what extent. Table 4.4 shows the amount and percentage of respondents' answers to different technological privacy tools questions. The most remarkable items were reported as follows: 37.72% of users always send private messages instead of posting on a Friend's wall and 31.74% do this most of the time. Half of the respondents (51.50%) reported that they have never gone offline on Facebook chat. The majority of the respondents (70.06%) said that they do not provide fake or inaccurate information to restrict other users. Nevertheless, almost half of the respondents (48.50%) are never worried about being embarrassed by wrong information others may post about them on Facebook, and also nearly half of the respondents (49.70%) are never concerned that others will see their profile.

	Never	Sometimes	About half the time	Most of the time	Always
Sending private messages	15 (8.98%)	25 (14.97%)	11 (6.59%)	53 (31.74%)	63 (37.72%)
Going offline on Facebook chat	86 (51.50%)	34 (20.36%)	14 (8.38%)	12 (7.19%)	21 (12.57%)
Excluding personal information	14 (8.38%)	35 (20.96%)	18 (10.78%)	52 (31.14%)	48 (28.74%)
Untagging from photos/videos	29 (17.37%)	81 (48.50%)	27 (16.17%)	21 (12.57%)	9 (5.39%)
Asking Friends to remove tags	80 (47.90%)	59 (35.33%)	17 (10.18%)	7 (4.19%)	4 (2.40%)
Deleting other users' posts	58 (34.73%)	67 (40.12%)	20 (11.98%)	14 (8.38%)	8 (4.79%)
Limiting access to profile	71 (42.51%)	41 (24.55%)	22 (13.17%)	14 (8.38%)	19 (11.38%)
Adding Friends into restricted list	87 (52.10%)	58 (34.73%)	15 (8.98%)	5 (2.99%)	2 (1.20%)
Providing fake or inaccurate info	117 (70.06%)	39 (23.35%)	7 (4.19%)	3 (1.80%)	1 (0.60%)
Worrying wrong info from others' posts	81 (48.50%)	55 (32.93%)	15 (8.98%)	14 (8.38%)	2 (1.20%)
Concerned about others seeing profile	83 (49.70%)	60 (35.93%)	11 (6.59%)	6 (3.59%)	7 (4.19%)

Table 4.4: The use of Technological Privacy Tools

Awareness of Social Privacy

The set of awareness of social privacy questions were aimed to show the extent of awareness of social privacy in terms of agreeing or disagreeing. Table 4.5 shows the amount and percentage of respondents' answers to different social privacy awareness questions. Neutral is expressed as "Neither agree nor disagree." The respondents answered the questions in a dispersed way, even so that the most remarkable items are that only 4.12% of respondents reported that they are worry about others scrutinizing their profile, while 30.00% reported they neither agree nor disagree. 43.53% of respondents said that they are uncomfortable with the level of exposure their Facebook content might bring.

	Strongly disagree	Somewhat disagree	Neutral	Somewhat agree	Strongly agree
Worrying others discovering my info	21 (12.35%)	39 (22.94%)	38 (22.35%)	56 (32.94%)	16 (9.41%)
Thinking who might be reading my post	27 (15.88%)	46 (27.06%)	35 (20.59%)	46 (27.06%)	16 (9.41%)
Thinking to whether people monitor my post	11 (6.47%)	43 (25.29%)	27 (15.88%)	58 (34.12%)	31 (18.24%)
Thinking about people creeping my profile	34 (20.00%)	43 (25.29%)	34 (20.00%)	51 (30.00%)	8 (4.71%)
Worrying others scrutinizing my profile	36 (21.18%)	35 (20.59%)	51 (30.00%)	41 (24.12%)	7 (4.12%)
Thinking who is reading my content	33 (19.41%)	46 (27.06%)	38 (22.35%)	41 (24.12%)	12 (7.06%)
Scrutinizing what info I post on my profile	19 (11.18%)	33 (19.41%)	43 (25.29%)	48 (28.24%)	27 (15.88%)
Not comfortable with level of exposure	7 (4.12%)	29 (17.06%)	35 (20.59%)	74 (43.53%)	25 (14.71%)
Worrying about people trying to creep me	17 (10.00%)	53 (31.18%)	32 (18.82%)	47 (27.65%)	21 (12.35%)
Thinking about who is monitoring me	18 (10.59%)	44 (25.88%)	31 (18.24%)	43 (25.29%)	34 (20.00%)

Table 4.5: Being Aware of Social Privacy

Figure 4.4 shows diverging bar plot of the standard scores (z-scores) of the awareness of social privacy with regard to the. The question items were computed and distributed regarding the value of items which are staying either above or below the average. The most prominent parts in this chart that the question item, not comfortable with the level of exposure, was above the average (z-score = 1.62); and, the item, worrying others scrutinizing my profile, was below the average (z-score = -1.24).

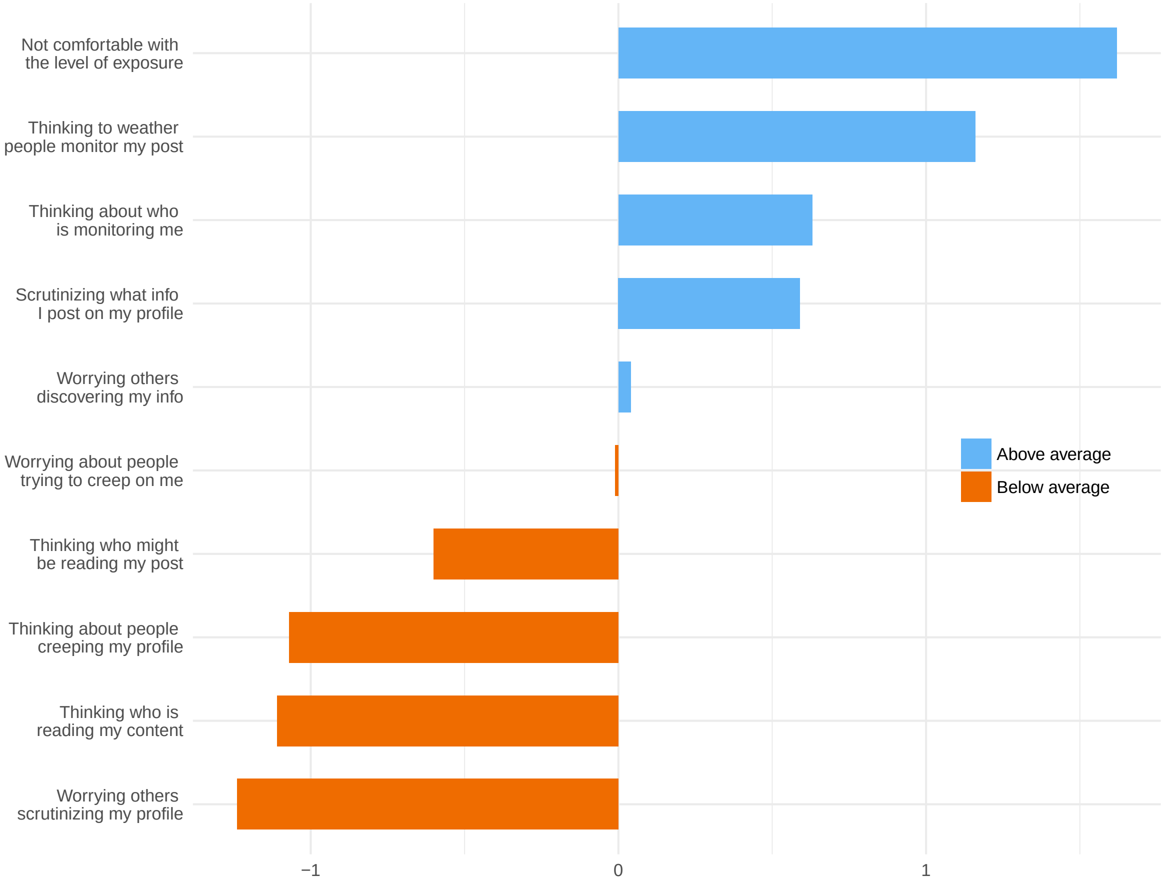


Figure 4.4: Diverging bar graph of standardized awareness of social privacy items

Negative Social Network Site Experiences

Negative SNS experiences reports the percentage of negative situations which respondents heard or experienced. In the data, 59.15% of respondents in the 19-25 age group said they heard some negative situation out of 74.3% of total respondents, while the other 25.7% did not. 68.90% of respondents in the 19-25 age group said they did not experience any negative situation out of 84.6% of total respondents, while 15.6% did. Overall, people heard about negative situations happening due to the disclosure of personal information online more than experiencing it.

4.1.3 Control variables

Figure 4.5 shows a scatter plot in which the dots accumulated on the right side of the scatter plot are more aware of social privacy. The standard deviation of the social privacy awareness was used as a measure of dispersion showing how the data spread out about the mean. The variables of Awareness of Social Privacy were calculated on the same scale as the sum of the total of construct variables, as that has better interpretability in the scatter plot chart. When we looked at the chart, the majority of dots accumulated in the upper-middle and left-middle places, showing these respondents showed a lesser awareness level.

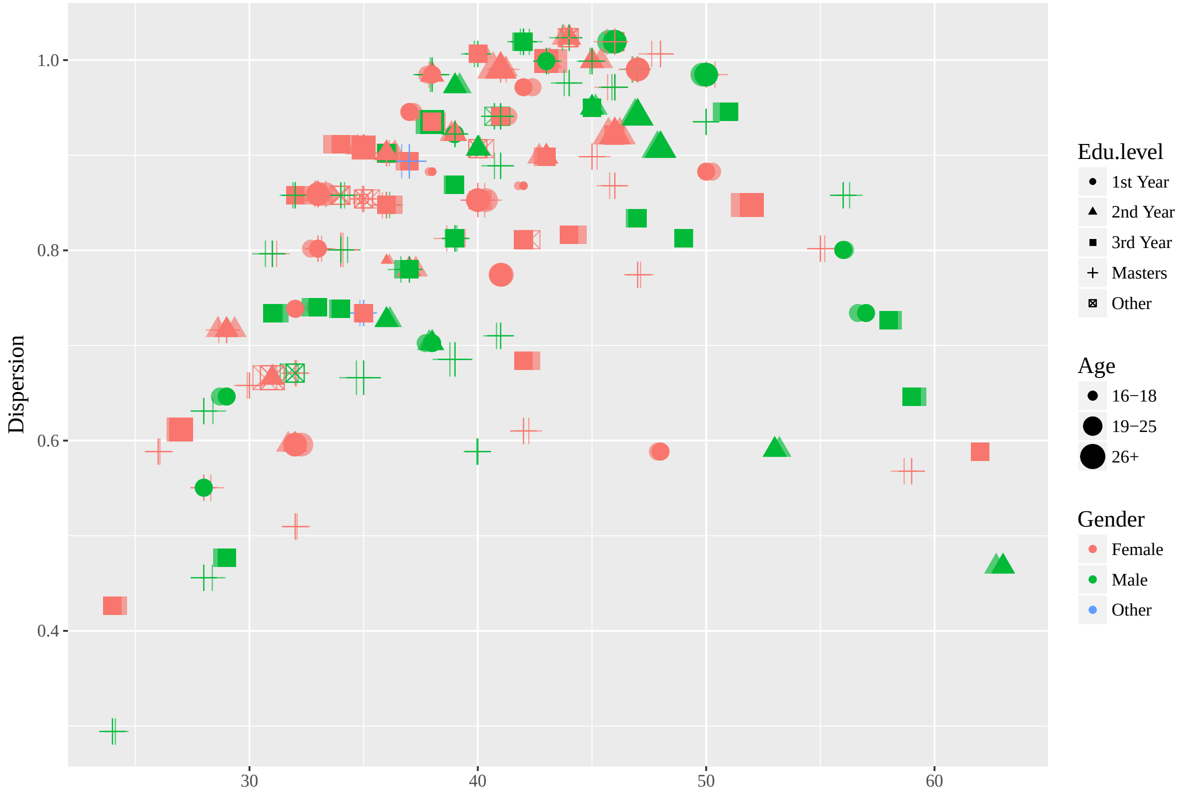


Figure 4.5: Scatter Plot of Total Level of Awareness of Social Privacy

In order to draw a conclusion about the relationship between demographic control variables and the dependent social privacy awareness variable, Pearson's chi square test was conducted to evaluate the likelihood of an observed difference between the sets (Field, 2009). Pearson's r was calculated to compare the frequency of the heard and experienced negative situations in demographic variables. No significant relationship was found in having social privacy awareness and gender $\left[X^2(44) = 113.27,\ p = .135\right]$. A significant relationship was found from social privacy awareness between age $\left[X^2(44) = 63.02,\ p = .003\right]$ and education level $\left[X^2(88) = 103.85,\ p = .034\right]$.

4.2 Validity and Reliability

Reliability and validity are two essential factors demonstrating the rigor of the research process and credibility of research findings (Roberts, 2006). Not only should the results of the study be significant, but also the rigor of the research (Heale & Twycross, 2015). In the survey analysis, reliability refers to the consistency and dependability allowing the results of analyses to recur in similar conditions (Neuman, 2014). Statistical tests are usually used to measure reliability. Internal consistency of the questionnaire questions can be found by an internal consistency test like Cronbach's alpha (α), which captures the standardized alpha based upon the correlations based on how coherent the scales used are. In this test, "the average of all correlations in every combination of split-halves is determined that instruments with questions that have more than two responses can be used (Heale & Twycross, 2015, p. 67)." The Cronbach's alpha (α) resulted between 0 and 1, and levels higher than .70 are accepted as a reliable analysis.

Assuring high quality results for the survey, the reliability and validity of the measurement scales were estimated. In order to examine the extent of internal consistency reliability to which the research instruments are related to other instruments (Heale & Twycross, 2015), Cronbach's alpha (α) is computed separately for the items in each construct. Table 4.6 illustrates the alpha value for each construct.

Construct	Cronbach's α
Intensity of Facebook Use	.387
Personal Information Disclosure	**.839**
Technological Privacy Tools	**.768**
Awareness of Social Privacy	.378
Negative SNS Experiences	.301
Note: Cronbach's alpha α value over .70 appears in bold.	

Table 4.6: Results of Cronbach's alpha (α) of the constructs

Personal Information Disclosure (α = .839) and Technological Privacy Tools (α = .716) signify reliable results. On the other side, the results that appeared in Intensity of Facebook Use, Awareness of Social Privacy, and Negative SNS Experiences were unreliable. Intensity of

Facebook Use (α = .387) was moderately reliable, but it had no strong reliability and correlation, and Awareness of Social Privacy (α = .378) and Negative SNS Experiences (α = .301) reveal unreliable results. These three unreliable constructs were re-examined to obtain a higher reliability result.

Table 4.7 demonstrates the item-total statistics of Intensity of Facebook Use, which consist of the correlations between items and the construct's total score. The 1st, 3rd, and 4th variables in the Intensity of Facebook Use construct had a negative correlation before they were reversed. After these items are reversed, if the low correlated items (6th, 7th, 8th, and 9th) led to a low alpha score, then they were removed from the construct. After all, the total scale of Cronbach's alpha (α) reaches the level of .482, which does not make the construct reliable enough. On that matter, if the construct was removed, then the variables were analyzed independently.

No	Item	Item-total score correlation	Cronbach's α with item dropped
1	How long have you been on Facebook? Since (2004-2017)	.609	.353
2	How many hours do you spend on checking your newsfeed on Facebook per day (average)?	.628	.484
3	How often do you go on Facebook?	.638	.316
4	How often do you share content on Facebook?	.510	.251
5	How many Friends do you have on Facebook?	.754	.540
	Level of connection with your Friends:		
6	Close friends	.378	.206
7	Acquaintances	.282	.132
8	Distant friends	.263	.151
9	People only met on Facebook	.217	.072

Table 4.7: Item-total statistics, Intensity of Facebook Use

Table 4.8 reveals the item-total statistics of Awareness of Social Privacy, which consisted of the correlations between items and construct's total score. When the negatively correlated items were reversed in 3rd, 9th, and 10th, the construct became reliable (α = .851). One of the reasons for having negatively correlated items was to use varied scales in the questions which did not increase the reliability of a construct. For instance, it was very obvious in the Intensity of Facebook Use construct where the construct variables implied different scales. Another reason is to have reverse coding in the items. The reverse coded items could lead to low correlations because they create confusion in the respondent's mind (Magazine, Williams, & Williams, 1996). Negative SNS Experiences was not recalled reliable as a construct (α = .301). So the construct was not employed because the constructs with less than three items are not suggested (Raubenheimer, 2004).

No	Item	Item-total score correlation	Cronbach's α with item dropped
1	Worrying others discovering my info	.541	.312
2	Thinking who might be reading my post	.583	.364
3	Thinking to whether people monitor my post	-.135	-.371
4	Thinking about people creeping my profile	.667	.472
5	Worrying others scrutinizing my profile	.603	.398
6	Thinking who is reading my content	.692	.501
7	Scrutinizing what info I post on my profile	.631	.435
8	Not comfortable with level of exposure	.483	.285
9	Worrying about people trying to creep me	-.104	-.334
10	Thinking about who is monitoring me	-.081	-.329

Table 4.8: Item-total statistics, Awareness of Social Privacy

After all, Negative SNS Experiences consisted of two variables, based on whether respondents had ever heard and experienced any negative situation happen because of disclosing personal information online, and which did not give enough reliability to set a construct. These variables were measured separately rather than forming them as a single construct. After this reliability test of the items, the validity test was taken.

Validity was discussed in terms of measurement validity, or how well the conceptual definition of the construct and empirical indicators fit together (Neuman, 2014). As shown in the Operationalization section, the questions measured the concepts intended to be measured because the questions of the survey were mainly constructed by the previous research, according valid indicators to the survey questionnaire. Therefore, the validity of the questionnaire could be tested with factor analysis (Emmons, 1984). On that account, a principal component factor analysis was performed to identify the validity implemented to check correlations and its structure between the variables.

A principal component analysis (PCA) was applied for all the items of the constructs with orthogonal rotation (varimax with Kaiser Normalization). The rotation sums of the squared loadings tried to maximize the variance of each factor. Firstly, Kaiser-Meyer-Olkin and Barlett's test of sphericity were performed to see whether it makes sense to do factor analysis on these variables (Field, 2009). The KMO that measures sampling adequacy indicated that the relationship among variables was high (KMO = .766), which was reasonable to continue the analysis. Barlett's test of sphericity was significant $\left[X^2(57) = 992.02,\ p < .05\right]$, indicating the assumption of equal variances was not valid. The number of principal components was assessed by Kaiser's criterion, which indicated the eigenvalues were greater than 1, and a parallel analysis ran with 100 simulations. Initially, only 6 components had eigenvalues greater than 1, which was the suggestion by Kaiser-Harris criterion. Table 4.9 illustrates the rotated factor loadings based upon the correlation matrix. The item variable names were shortened with their unique variable numbers, such as I (Personal Information Disclosure), T (Technological Privacy Tools), and A (Awareness of Social Privacy), in order to display an efficient table.

	Rotated Components					
Item	1	2	3	4	5	6
I1					.102	**.543**
I2			.185	.141	.123	.356
I3	.160				.102	**.659**
I4					**.765**	.196
I5			.112		**.734**	.148
I6		.271		.141	**.540**	
I7	.200	**.722**				.155
I8		.493		.227	.444	
I9	.116	**.546**		.226	-.219	.263
I10	.173	.122	.132	.232		**.561**
I11		**.571**		.148	.367	-.132
I12		**.549**	.133	.122	.412	
I13	.113	**.558**			.223	
I14	.109	**.739**	.102			
I15	.137	.355		.127		.354
I16		-.108		**.686**		.330
I17			.110	**.720**	-.100	.291
I18		.251		**.759**		
I19		.443		**.547**	.323	-.166
I20		.173		**.721**		
I21		.409	-.113	.406	.204	-.168
I22		.266		**.554**		.362
T1	.211		-.117	.434	.207	-.181
T2	.147	.295	.283			
T3	.138	.162	.264	.330	.206	-.123
T4		.240	**.603**			.116
T5		.126	**.663**			
T6	.138		**.728**		.131	
T7		-.218	**.628**		.145	
T8		-.194	**.609**			.127
T9	.208		**.589**	.126	-.141	
T10	.150		**.629**		-.176	
T11	.520		**.539**		-.123	
A1	**.593**	.212				.179
A2	**.737**			-.101		
A3	**.573**	.122	.165			
A4	**.787**	.218				.107
A5	**.798**		.112			.106
A6	**.742**		.217			
A7	.457		.265	.209		-.305
A8	.332		.166	.212		-.230
A9	**.639**					.221
A10	**.632**	.100	.135			.204
Eigenvalues	4.784	3.736	3.661	3.579	2.451	2.262
Proportion S^2	.111	.087	.085	.083	.057	.053
Cumulative S^2	.111	.198	.283	.367	.424	.476

Note: A factor loading over .50 appears in bold.

Table 4.9: Rotated factor loadings based upon correlation matrix

Since five or more factor loadings with the value of .50 or better were considered suitable to imply a solid factor (Osborne & Costello, 2005, p. 5), the factor loadings over .50 were chosen to create a construct. When there were more than two component values over .50, the higher factor loading between them was chosen. The 1st, 2nd, 3rd, and 4th components showed five or more factor loadings higher than .50. The Awareness of Social Privacy and Technological Privacy Tools were reconstructed according to the 1st and 3rd components' factor loadings respectively. The variables less than .50 were removed from the construct. The Personal Information Disclosure seemed to divide into two components, which were reconstructed from the 2nd component and the 4th component. "Basic information disclosure" was constructed from the loaded items of the 2nd component whose variables are such as "Family members", and "Partner's name"; "Appearance information disclosure" was constructed from the loaded items of the 4th component whose variables are such as "Photos of you", and "Places you visit." After all, the constructs new KMO variable was .798 which was reasonable to perform factor analysis. Four components having eigenvalues greater than 1 were extracted. Table 1B and Table 2B in Appendix B show the new reliability and validity constructs. The factor analysis changed the constructs and measures of the constructs: Awareness of Social Privacy, Technological Privacy Tools, Basic Personal Information Disclosure, and Appearance Personal Information Disclosure.

4.2.1 Correlation Matrix

Subsequently, a correlation matrix was performed to display the initial overview of the relations between variables with the statistical significance of the variables checked by Pearson's correlation coefficients. Table 4.10 displays the correlation matrix of independent and dependent variables with Pearson's correlation coefficient significance.

		Correlations								
		1	2	3	4	5	6	7	8	9
1	Basic	—								
2	Appear	.394**	—							
3	Techno	.291**	.142	—						
4	Aware	.157*	.114	.361**	—					
5	Intensity	-.161*	-.109	-.091	-.051	—				
6	Negative	-.213**	-.224**	-.205**	-.208**	.061	—			
7	Gender	.074	.045*	.047	.048	.006	.025	—		
8	Age	-.021	.031	.141	.084	.041	.061	-.044	—	
9	EduLevel	.109	.071	.092	.054	-.024	.081	-.052	.531**	—

N=169
* Correlation is significant at the 0.05 level
** Correlation is significant at the 0.01 level

Table 4.10: Pearson's correlation of independent and dependent variables.

Some interesting correlation relations were observed that awareness of social privacy and previous negative SNS experiences were negatively related [$r(169)$ = -.208, $p < .01$], which may indicate that this reduces social privacy awareness. Nonetheless, the other variables, such as intensity of Facebook use, age, and gender were positively related to privacy awareness, inferring higher social privacy awareness as a result. However, these variables did not significantly correlate with awareness of social privacy, except technological privacy tools showing the highest correlation value among others. The use of technological privacy tools had a positive correlation with awareness of social privacy [$r(169) = .361$, $p < .01$], which means that the use increases as social privacy awareness increases. The dependent variable of basic personal information disclosure showed positive correlations with all the variables except intensity of Facebook use [$r(169)$ = -.161, $p < .05$], previous negative SNS experiences [$r(169)$ = -.213, $p < .01$], and age [$r(169)$ = -.021]. In particular, age showed a negative correlation because the respondents mainly belong to the 16 to 25 years age group. Among the positive correlations, gender showed the lowest value [$r(169) = .074$] among the others caused by females being less likely to share basic kinds of information in comparison to male. Surprisingly, the previous negative SNS experiences showed the negative correlation with social privacy awareness [$r(169)$ = -.208, $p < .01$] that implies users who are exposed to negative SNS experiences (either heard or experienced) have lower social privacy awareness. Besides, it is found that previous negative SNS experiences were negatively correlated with basic and appearance personal information disclosure, technological privacy tools, and education level.

Overall, the correlations between these variables implied that the measurements and relationships were highly varied. Several regression analyses were conducted to look at possible inferred causal relationships between these variables. Following that, the average of both variables within constructions was calculated before conducting the regression analysis as variables in the constructs stipulated associations. The results of the reliability and factor analysis led to the creation of new variables. The item scores of the constructs were recalculated as the average, first added up and then divided by the number of items, for each respondent.

The other variables were standardized as mean-centered, meaning both response and predictor values were centered by their means. Moreover, standardizing the variables was utilized to reduce misleading results due to multicollinearity, which led to large confidence intervals and a difficult coefficients interpretation (Kabacoff, 2015). Following that, the interaction variables, which help understand the sole effect between a response and predictor variable, were also standardized by centering values by their means.

4.3 Hypotheses Tests

A set of Multiple Linear Regression (OLS) analyses were performed in order to test the hypotheses proposed in the Theoretical Framework, mainly to explore the dependent variable, social privacy awareness, with other proposed independent variables. In addition, the demographic data (gender, age, education level) were used as control variables in the regression models in order to see whether the predictor values are significant, and whether the predictor values have unique variance. The nationality variable was not taken into consideration because of high variety, and the department variable was not included because it was collected as an open-ended question. The categorical gender variable is recoded as a dummy variable (0 = Male and 1 = Female) for the subsequent regression analyses. "Other" was omitted due to low data in the sample population. Considering that regression analyses were done with social privacy awareness, several independent variables were constructed as reliable and valid in the previous analyses.

Before performing the regression models, the assumptions of multiple linear regression, also regression diagnostics, were checked in terms of linearity, normality, (multi)collinearity, no outliers, and homoscedasticity that these assumptions should not be violated[3]. First, linearity was checked to assume there was no relation between residuals and fitted values, and it was illustrated in the Residuals vs Fitted plot of dependent and independent variables that shows if the residuals indicate any non-linear patterns. Second, the multivariate normality distribution is checked by the normal probability plot of standardized residuals (shown in Normal Q-Q plot). The points of the graph should accumulate on the 45-degree line, otherwise the normality assumption was more likely to be violated (Kabacoff, 2015). Besides this, an extra Shapiro-Wilk test is performed to check all the departures from normality, including residuals. Third, multicollinearity was observed to ensure that no or little independent variables were independent of each other. It was calculated with the variance inflation factor (VIF), which generally has to be smaller than 5 (Field, Miles, & Field, 2012). Fourth, outliers can be seen in the Residuals vs Leverage plot which also identifies high-leverage points and individual observations (Kabacoff, 2015). The graph indicated outliers; however, not all outliers should be considered as influential to the result of the regression analysis, except the cases outside of Cook's distance. Finally, homoscedasticity was tested by plotting the square root of standardized residuals against fitted values in the Scale-Location plot.

4.3.1 Personal Information Disclosure

Two regression models were used to examine the relation between personal information disclosure and social privacy awareness (H_1). In the first model, the basic personal information disclosure on social privacy awareness was tested. In the second model, the appearance personal information disclosure on social privacy awareness was tested.

Before the regression analysis, the regression diagnostics were checked. Figure 4.6 shows the

[3] The assumptions tests are also visualized in the plots as the numerical calculations without graphs can be tricky, considering the famous Anscombe's quartet (O'Connor, 2014).

diagnostic plots for the regression analysis of basic personal information on social privacy awareness. The first plot (Residuals vs Fitted) shows a fairly linear relationship that does not indicate any distinctive pattern that assumes the linearity assumption is violated. Second, the Normal Q-Q plot shows the normality assumption was met as the points forming the line were close to being straight. Third, the Scale-Location plot implies that homoscedasticity was not met as the points on the graph did not equally spread along the horizontal line. Fourth, the last plot, Residuals vs Leverage, showed that there were some influential cases pointed outside of Cook's distance line. The analysis of multicollinearity was completed ($VIF = 1.02$), and collinearity was not violated by this test. Additionally, for the normality distribution of residuals, the Shapiro-Wilk test was performed, finding that the residuals did deviate from normality ($p < .001$).

To keep the results section neat, the regression diagnostics plot for appearance personal information was not reported here because the graphs show a high similarity to the basic one. The explanation for regression diagnostics was also similar. Moreover, Shapiro-Wilk test analyzing the normality distribution of residuals was performed that the residuals also did deviate from normality ($p < .001$).

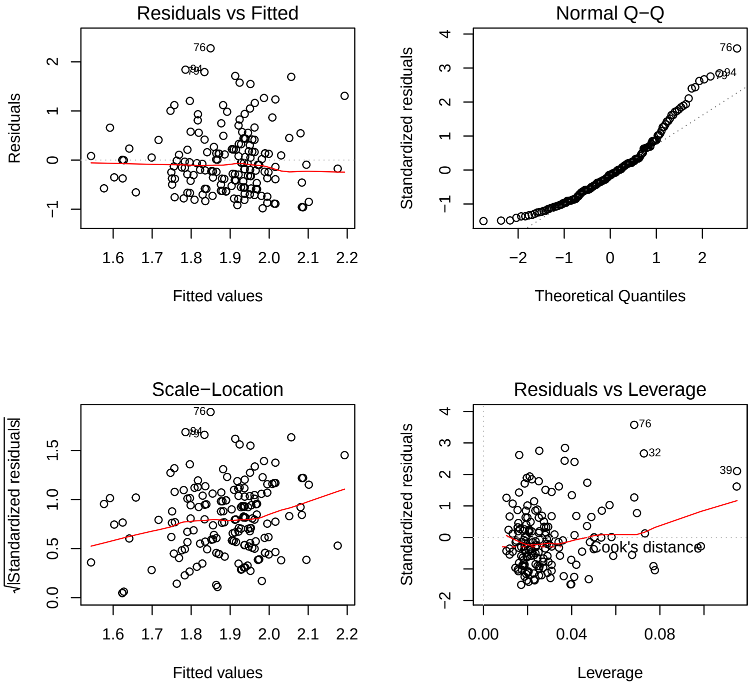


Figure 4.6: Diagnostic plots for the regression of basic information on privacy awareness

The regression results of regression coefficients and standard error from the first and the second model were reported in Table 4.11. The regression models of the level of social privacy awareness, as the dependent variable, and the basic information disclosure, $F(4,\ 164) = 1.218$, $p = 0.30$, and appearance personal information disclosure, $F(4,\ 164) = 1.218$, $p = 0.30$, as independent variables, were found no significant along with the control variables, the age, gender, education level.

	Model 1		Model 2	
	b^*	SE	b^*	SE
(Intercept)	1.3005***	.288	1.5328***	.240
Control variables				
Gender	.003	.104	.010	.105
Age	.177	.018	.015	.018
Education Level	-.009	.049	-.007	.049
Independent variables				
Basic Personal Information	.196	.100	—	—
Appearance Personal Information	—	—	0.124	0.09
R^2	.288		.016	
F	1.218		0.699	
ΔR^2			.272	
ΔF			0.519	
Significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$.				

Table 4.11: Regression results for basic and appearance personal information disclosure

Previous studies have depicted a strong association between personal information disclosure and social privacy awareness, as well as how more aware users are less likely to share on Facebook (Acquisti & Gross, 2006). On the other hand, there was a general trend that users tend to share their personal information online despite awareness. The users could decide to share because of the trade-off of obtaining satisfaction or just drawing attention to build their social status (Debatin et al., 2009; Gross & Acquisti, 2005). However, this could not be directly associated with social privacy awareness, as they could still be aware but could opt for the trade-off. A further measurement is needed to understand this distinction.

When considering the personal information disclosure on an individual basis, some personal information items (such as home address) were positively correlated with social privacy awareness. The first hypothesis, which expected an association between the amount of disclosed information as a whole (not based on individual items) of certain types and social privacy awareness, indicates that the level and kinds of personal information disclosure do not increase or decrease social privacy awareness. The relationship between the two was not supported in the reliability and validity test; splitting personal information into two categories returned reliable and valid constructs, which were named basic and appearance personal information. However, these two constructs did not show a significant result with the dependent variable and with other control variables. Besides that, the regression assumptions for these constructs were not met, which violated the regression results in normality and homoscedasticity terms. Although the hypothesis was not significant, the comparison of the means for the different items of personal information disclosure with the varying levels of sharing, not sharing, and sharing but inaccurate gave interesting results in the bar graph. Thus, it seems that there was a pattern in the relation between social privacy awareness and individual personal information items, though this pattern could not be statistically proven. When the regression analysis was run between these two, it did not return a statistically significant result, as social privacy awareness was constructed as an average mean. The ample data and variables in social privacy awareness made this comparison between the two incomprehensible. All in all, this result showed us that this hypothesis for personal information disclosure for the two constructs was not supported, but people have concerns about disclosing certain personal information online.

4.3.2 Intensity of Facebook Use

A two-way factorial ANOVA was performed to examine the effects of social privacy awareness in connection with the effects in kinds of intensity of Facebook use ($H_{2\text{a}}$). Table 4.12 illustrates that the dependent variable of the level of social privacy awareness was performed on the dependent variables, and only the frequency of going on Facebook returned a statistical significance $\big[F(4,\ 169) = 3.706,\ p = 0.006\ \big]$. Nevertheless, the significance value generated in a two-way analysis of variance does not tell us where this effect happens. Since the frequency of going on is between six levels, determining which conditions are significantly different from other conditions requires conducting and reporting the results of a post-hoc test, which compares the significance of each condition with all other conditions (Field et al., 2012). A post-hoc comparison using Tukey HSD (Tukey's Honest Significance) test was conducted (with the confidence level of 0.95) on all possible family-wise contrasts and multiple comparisons of means to see the differences between means of the specified variables. The dependent variable, awareness of social privacy, was mean-centered standardized. Since Tukey HSD requires categorical variables for the test, the continuous dependent variable, awareness of social privacy, was categorized in the terms of number ranges in four degrees, like 1 to 2, 2 to 3, 3 to 4, and 4 to 5 (out of a five point Likert scale ranging from strongly agree to strongly disagree). Tukey HSD showed that the groups between the "3 to 4" level of social privacy awareness and "once a day" going on Facebook differed significantly at $p < .05$.

	Sum of Squares	*df*	*MS*	*F*	*p*
Year of being on FB	0.04	1	0.04	0.093	0.761
Hours of checking news feed	1.17	3	0.39	0.891	0.447
The frequency of going on FB	6.09	4	1.52	3.706	0.006**
The frequency of sharing content	3.93	5	0.785	1.842	0.107
The number of Friends	1.75	4	0.43	1.005	0.409
Level of connection with Friends					
Close friends	1.42	3	0.474	1.086	0.356
Acquaintances	0.82	3	0.274	0.625	0.601
Distant friends	1.10	3	0.367	0.869	0.474
People only met on FB	1.04	3	0.346	0.789	0.502
Awareness of Social Privacy (DV)					
Significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$.					

Table 4.12: Results of the two-way analysis of variance on awareness of social privacy

A two-way factorial ANOVA was performed to examine the effects of technological privacy tools in connection with the effects of kinds of intensity of Facebook use ($H_{2\mathrm{b}}$). Table 4.13 illustrates that the dependent variable of the level of technological privacy tools was performed on the independent variables, and only the frequency of sharing content returned a statistical significance $\left[F(5,\, 169) = 2.862,\, p = 0.02\ \right]$. The dependent variable, technological privacy tools, was mean-centered standardized. Since Tukey HSD requires categorical variables for the test, the continuous dependent variable, technological privacy tools, was categorized in the terms of number ranges in four degrees, like 1 to 2, 2 to 3, 3 to 4, and 4 to 5 (out of five point Likert scale ranging from never to always). A further Tukey HSD comparison of means showed that the groups between "2 to 3" of the use of technological privacy tools and "never" of the frequency of sharing content differed significantly at $p < .05$.

	Sum of Squares	*df*	*MS*	*F*	*p*
Year of being on FB	0.03	1	0.03	0.04	0.842
Hours of checking news feed	2.63	3	0.875	1.12	0.343
The frequency of going on FB	2.76	4	0.690	0.879	0.478
The frequency of sharing content	9.73	5	1.946	2.862	0.02*
The number of Friends	2.17	4	0.543	0.688	0.601
Level of connection with Friends					
Close friends	2.22	3	0.737	0.941	0.422
Acquaintances	2.26	3	0.754	0.962	0.412
Distant friends	0.46	3	0.153	0.193	0.901
People only met on FB	1.79	3	0.595	0.757	0.521
Technological Privacy Tools (DV)					
Significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$.					

Table 4.13: Results of the two-way analysis of variance on technological privacy tools

Intensity of Facebook use was assessed in order to see the associations with social privacy awareness and technological privacy tools. The previous studies concluded a greater correlation between intensity of Facebook use and the use of technological privacy tools: that the users who regularly use Facebook have an increased use of technological privacy tools (Boyd & Hargittai, 2010), and active users are more aware of social privacy (Litt, 2013). Intensity of Facebook use was not built as a construct as it did not come up with reliable results, so the effect between these associations was assessed individually, and the results were reported with the ones returning statistical significance. The results illustrated that the frequency of users going on Facebook has shown a significance with social privacy awareness. The users who were going on Facebook more often were more aware of social privacy. However, this result did not align with any theoretical framework. The result was statistically significant; however further analyses or data might be needed to come up with a new, consistent result. On the other side, the frequency of sharing content is associated with the use of technological privacy tools. A possible explanation for this is that the users who share more personal information become more competent with the use of these settings. As social privacy awareness was related to the knowledge of using of technological privacy tools (Litt, 2013), the frequency of sharing content could grow awareness. With that explanation, users were more exposed to the practices of sharing and were more likely to pay attention to unwanted audiences who may be seeing their personal their personal information (Young & Quan-Haase, 2013). Nevertheless, the population sample could be the reason for the increment between these two associations. Thus, these hypotheses were partially supported.

4.3.3 Technological Privacy Tools

A regression model was used to examine the relation between the use of technological privacy tools and social privacy awareness (H_3). Before the regression analysis, the regression diagnostics were checked. Figure 4.7 shows the diagnostic plots for the regression analysis of basic personal information on social privacy awareness. The first plot (Residuals vs Fitted) shows a fairly linear relationship that did not indicate any distinctive pattern that assumes the linearity assumption was violated. Second, the Normal Q-Q plot showed the normality assumption was met as the points forming the line were close to being straight. Third, the Scale-Location plot implied that homoscedasticity was not met as the points on the graph did not equally spread along the horizontal line. Fourth, the last plot, Residuals vs Leverage, showed that there were some influential cases pointed outside of Cook's distance line. Also it shows a few outliers. Furthermore, The analysis of multicollinearity was completed ($VIF =$ 1.14), and collinearity was not violated by this test. Additionally, for the normality distribution of residuals, the Shapiro-Wilk test was performed, and found that the residuals did deviate from normality ($p < .001$).

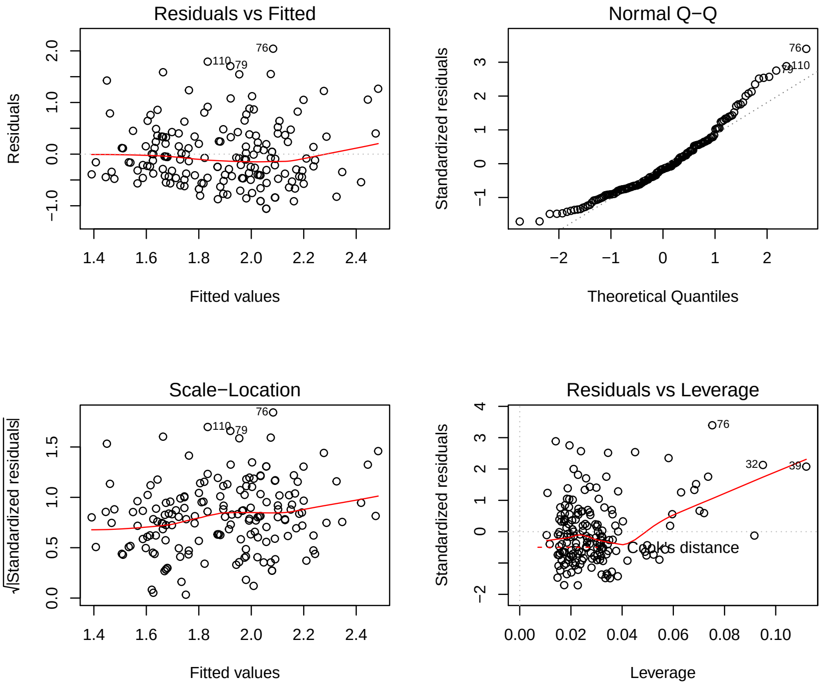


Figure 4.7: Diagnostic plots for the regression of technological tools on privacy awareness

The regression results of regression coefficients and standard error from the model were reported in Table 4.14. The regression models of the level of social privacy awareness, as the dependent variable, and the technological privacy tools, as independent variable, was found significant, $F(4, 164) = 5.956$, $p < .001$. This regression model was thus useful for predicting the social privacy awareness; however, no other variables in the control group were found to be significant.

	$b*$	SE
(Intercept)	1.1128***	.206
Technological Privacy Tools	.262***	.055
Control variables		
Gender	.011	.099
Age	.005	.017
Education Level	-.006	.004
R^2	.126	
F	5.956***	
Significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$.		

Table 4.14: Regression results for technological privacy tools

The use of technological privacy tools on social privacy awareness did not return a significant effect. Supposedly, the construction was merged and omitted the crucial analysis of the items situated in the technological privacy tools. The results of the effect of technological privacy tools on social privacy awareness are in contrast to the previous study in which they are strongly associated Acquisti and Gross (2006), and users should be more aware of privacy if their use of privacy tools was high (Young & Quan-Haase, 2013). On the other hand, it was seen in the association of intensity of Facebook use and privacy tools that social privacy awareness could have several dimensions, so a mean-centered construct could be weak for a holistic result, meaning the items of constructs should have been analyzed independently. However, it was not quite possible because the social privacy awareness was also mean-centered and both of these constructs had multiple items that would cause a complicated statistical analysis. The use of privacy tools and social privacy awareness resulted as in significant since they are constructed as a whole, so the problem could come from two points: first, they should not be mean-centered constructed, and second the hypothesis should have been tested with another variable, such as a negative past experience. The use of technological privacy tools could give better results if they are associated with consequences (Christofides et al., 2009). The following hypothesis (H_4) was tested by Welch's two sample independent t-test to compare the means between the dependent variable, social privacy awareness, and the independent variable, changing default recommended privacy settings. A further Levene's test was not applied for the equality of variances as the Welch's t-test is very robust (Kabacoff, 2015). A significant difference was found in social privacy awareness in the conditions of changing default privacy settings. On average, "Yes" scored (M = 1.82) lesser than "No" (M = 1.90) in the heard group (1st plot), $t(86)$ = -2.39, $p < .001$. This result suggests that changing default recommended settings do have an effect on social privacy awareness; and it is assumed that when users change default privacy settings, they become more aware of social privacy.

The results show that the change in default recommended privacy settings on social privacy awareness was significant, that if the users have changed their settings, an option which is offered in the beginning of Facebook registration, they are more aware. A previous study reported that many users do not touch the default settings that Facebook offers, which usually are set to a open and public — for most information, so that the knowledge of the option to change default settings results in social privacy awareness (Debatin et al., 2009).

4.3.4 Negative Social Network Site Experiences

The aforementioned reliability analysis showed that the variables are not reliable enough to build a construct, because the variables were measured independently from each other. Therefore, several Welch's two sample independent t-tests were run to reveal the relationship between the dependent variable, social privacy awareness, and the independent variables, consisting of the variables. The means of these variables were compared to measure negative SNS experiences that are heard negative situation (H_5) and experienced negative situation

(H_6).

Figure 4.8 illustrates the notched box plots of sample mean estimates for the negative SNS experiences on social privacy awareness. First, a strong significant difference was found in social privacy awareness in the conditions of heard and not heard. On average, "Yes" scored (M = 1.95) higher than "No" (M = 1.70) in the heard group (1st plot), $t(86)$ = -2.39, $p < .001$. This result suggests that having heard about a negative experience has an effect on social privacy awareness; when users hear about negative experiences, they become more aware of social privacy. Second, a significant difference in social privacy awareness was found in the conditions of experienced and not experienced. On average, "Yes" scored (M = 2.29) significantly higher than "No" (M = 1.81) in the heard group (2nd plot), $t(33)$ = -3.03, p = .004 . It is implying that having experienced a negative situation has an effect on social privacy awareness; when users have negative experiences, they become more aware of social privacy.

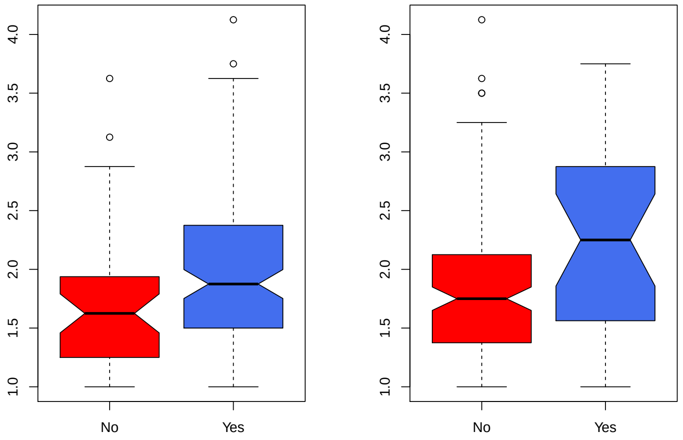


Figure 4.8: Box plots of heard and experienced negative SNS from a two sample t-test

Social privacy awareness was assessed with respect to previous negative situations, both heard and experienced, which both returned statistically significant results. In their study, Debatin et al. (2009) and Young and Quan-Haase (2013) found that the majority of the students make a change in their privacy settings, or start using privacy tools, when they have a negative experience, which could be either heard or experienced. We assumed that this change was one part of the increase in social privacy awareness; that is to say that the students who had these experiences were more likely to be more aware of social privacy than those who did not. The users who have had negative events in the past take some precautions to avoid similar events that may happen in the future. These precautions show that they become more aware of their personal information and social surveillance practices, as the negative experiences play an educational role. Having a negative, unpleasant experience in the past was a strong factor to social privacy awareness (Horváth et al., 2014). The student users who have had negative experiences in the past tend to restrict their disclosure and increase the use of technological tools (Christofides et al., 2009). These two actions enhance the level of social privacy awareness. Thus, this hypothesis was confirmed, and it was in line with our expectations. Litt (2013) reported that some users take privacy cautions after they receive unpleasant messages from other users. These privacy cautions can be various, like decreasing the amount and type of personal information disclosure, changing the visibility of posts, or blocking people with technological privacy tools.

4.3.5 Summary of Hypothesis Testing

Hypotheses		
H_1	Personal information disclosure ⟶ Social privacy awareness	Not Supported
$H_{2\mathrm{a}}$	Intensity of Facebook use ⟶ Social privacy awareness	Partially Supported
$H_{2\mathrm{b}}$	Intensity of Facebook use ⟶ Technological privacy tools	Partially Supported
H_3	Technological privacy tools ⟶ Social privacy awareness	Not Supported
H_4	Default privacy settings⟶ Social privacy awareness	Supported
H_5	Heard negative situation ⟶ Social privacy awareness	Supported
H_6	Experienced negative situation ⟶ Social privacy awareness	Supported

Table 4.15: Summary of hypotheses test results

5 Conclusion

This chapter offers a conclusion of the findings of the study, encountered limitations, and recommendations for possible future research. This thesis has studied the extent of social privacy awareness among university students in the Netherlands. Social privacy awareness is constituted in terms of information disclosure, social surveillance and visibility strategies. This thesis employed a survey research in order to measure the social privacy phenomenon in accordance with various variables. The level of personal information disclosure, intensity of Facebook use, the use of technological privacy tools, and negative social network sites experiences are quantitatively analyzed. The multiple regression analyses, ANOVA, and Welch's *t*-test are used to examine the relationships regarding student privacy awareness. Also, gender, age, and level of education are included as control variables in the analysis. The findings reported that social privacy awareness has a strong association with negative social network site experiences, and a less strong association has been found with the use of technological privacy tools and intensity of Facebook use. Based on the reliability analysis, personal information disclosure has led to the distinction of two constructed groups: basic personal information, which consists of the most basic information of a user such as their full name, and appearance personal information, which is based upon what users show to others such as a profile picture. The findings of this study clearly demonstrate that users have a high level of social privacy awareness when they changed their default privacy settings and heard or experienced a negative situation. Three questions were assessed: to what extent the university students share their personal information, what technological privacy tools they employ, and what their general attitude is towards social surveillance. These assessments give a clear understanding of the extent of awareness of social privacy among university students in the Netherlands. The intensity of Facebook use on technological privacy tools and social privacy awareness were partially supported. Overall, the study shows that Facebook users seem to disclose less information about themselves and consistently use technological privacy tools, but these variables should not be considered on a scale.

Limitations and Future Research

This thesis study shows that Facebook users seem to disclose less information about themselves and consistently use technological privacy tools. There are various limitations identified in this thesis study. Although the results did not show significant privacy concerns in the way they were implied, a further measurement will be necessary to reveal the patterns between these associations.

As shown in the descriptive results, the majority of Facebook users seem aware of social privacy. However, the reasons for disclosing information and their attention to their own visibility were not measured. Besides that, the number of respondents who seem aware of social privacy but disregard the privacy issue because of trade-off advantages is unknown. A further measurement may be needed to understand the distinction between privacy awareness and trade-offs, which means that sharing has more advantages than disadvantages in the case of privacy reasons. Personal information disclosure should be analyzed and distinctive from social privacy, as users are aware of privacy but share due to reasons regarding the desire of publicity or validation from social ties (Trottier, 2012b), or they are just not aware of social privacy.

As the population focused on university students (Millennials), who are mainly literate and educated, there can be a sample bias to make generalizable results for the other students. Moreover, the research is focused on students who had Facebook accounts. Knowing the behavior of non-users and making a comparison between them could lead to interesting results. This may require further explanation and maybe a comparison study that shows how social privacy awareness would be different for the people who do not use Facebook or social network sites in general.

Another kind of sample bias was that the majority of the respondents were female. Also, the respondents were mainly coming from natural sciences departments. Additionally, a larger number of respondents would decrease the bias caused by the sample size. Another limitation is that the patterns of social privacy awareness could be measured in a less vast and more detailed way that would better fit the inferential statistics. Avoiding this could be overcome by conducting a small plot survey, eliminating the questions that do not work, and testing the relationship between variables. Additionally, Facebook users have asserted that they are aware of the presence of privacy settings and that they know how to use them. As we cannot be sure about the reliability of that claim, maybe a qualitative study can uncover that issue.

The survey was not quite effective for asking what kind of information is visible to whom on Facebook; instead, a content analysis parsing the respondents' public Facebook profiles in comparison with survey results would give more robust results. The common risk in the questions revealing personal information disclosure is the people, who do not know or simply forget what they share and to whom it is visible. Furthermore, this approach harms the anonymity of a survey study; a case study, therefore, would be more appropriate for this analysis.

Future research will need to examine social privacy awareness in relationship with these variables and in consideration with the limitations this thesis has encountered. Adding a new theoretical discussion associating social privacy with, for example, the level of trust for these variables or the issue of risk and trade-offs has great potential to enhance the research. A new study with more consistent measures will be useful.

References

Acquisti, A. & Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In G. Danezis & P. Golle (Eds.), *Privacy enhancing technologies. pet 2006.* (Vol. 4258, pp. 36–58). Berlin, Heidelberg: Springer.

Albrechtslund, A. (2008). Online social networking as participatory surveillance. *First Monday*, *13*(3). doi:10.5210/fm.v13i3.2142

Andrejevic, M. (2005). The work of watching one another: Lateral surveillance, risk, and governance. *Surveillance & Society "People Watching People" (ed. Wood)*, *2*(4), 479–497. Retrieved from http://www.surveillance-and-society.org

Bakhshi, S., Shamma, D. A., & Gilbert, E. (2014). Faces engage us: Photos with faces attract more likes and comments on Instagram. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 965–974). CHI '14. New York, NY, USA: ACM. doi:10.1145/2556288.2557403

Berger, P. L. & Luckmann, T. (1991). *The social construction of reality: A treatise in the sociology of knowledge.* London: Penguin UK.

Blank, G. & Reisdorf, B. C. (2012). The participatory web: A user perspective on Web 2.0. *Information, Communication & Society*, *15*(4), 537–554.

Bloustein, E. J. (1964). Privacy as an aspect of human dignity: An answer to Dean Prosser. *NYUL Rev. 39*(6), 962–1007.

Boyd, D. & Ellison, N. B. (2008). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. doi:10.1111/j.1083-6101.2007.00393.x

Boyd, D. & Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday*, *15*(8), 1–12. doi:10.5210/fm.v15i8.3086

Boyd, D. & Marwick, A. (2011). Social privacy in networked publics : Teens' attitudes, practices, and strategies. *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, September 2011*, 1–29. doi:10.1037/0003-066X.63.2.111

Brighenti, A. (2007). Visibility: A category for the social sciences. *Current Sociology*, *55*(3), 323–342. doi:10.1177/0011392107076079

Buttle, F. A. (1996). Servqual: review, critique, research agenda. *European Journal of Marketing*, *30*(1), 8–32. doi:10.1108/03090569610105762

Chad, B. (2016). Social screening: What hiring managers look for on social media. Retrieved from http://www.businessnewsdaily.com/2377-social-media-hiring.html

Child, J. T. & Starcher, S. C. (2016). Fuzzy Facebook privacy boundaries: Exploring mediated lurking, vague-booking, and Facebook privacy management. *Computers in Human Behavior*, *54*, 483–490. doi:10.1016/j.chb.2015.08.035

Christofides, E., Muise, A., & Desmarais, S. (2009). Information disclosure and control on Facebook: Are they two sides of the same coin or two different processes? *Cyberpsychology & Behavior*, *12*(3), 341–345. doi:10.1089/cpb.2008.0226

Clarke, R. (1999). Internet privacy concerns confirm the case for intervention. *Commun. ACM*, *42*(2), 60–67. doi:10.1145/293411.293475

Cohen, J. E. (2008). Privacy, visibility, transparency, and exposure. *Georgetown University Law Center*, *75*(1), 181–201.

Das, S. & Kramer, A. (2013). Self-censorship on Facebook. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, 120–127. doi:10.1007/b104039

De Leeuw, E. (2008). Choosing the method of data collection. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 113–135). New York, NY, USA: Taylor & Francis Group.

Debatin, B., Lovejoy, J. P., Horn, A. K., & Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication*, *15*(1), 83–108. doi:10.1111/j.1083-6101.2009.01494.x

Dinev, T. & Hart, P. (2005). Internet privacy concerns and social awareness as determinants of intention to transact. *International Journal of Electronic Commerce*, *10*(2), 7–29. doi:10.2753/JEC1086-4415100201

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of facebook "friends": Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168. doi:10.1111/j.1083-6101.2007.00367.x

Emmons, R. A. (1984). Factor analysis and construct validity of the narcissistic personality inventory. *Journal of Personality Assessment*, *48*(3), 291–300.

Facebook. (2016). Facebook company info. Retrieved from https://newsroom.fb.com/company-info/

Facebook. (2017). Basic privacy settings tools. Retrieved from https://www.facebook.com/help/211513702214269

Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed.). London: SAGE Publications.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London: SAGE Publications.

Fuchs, C. & Trottier, D. (2015). Towards a theoretical model of social media surveillance in contemporary society. *Communications*, *40*(1), 113–135. doi:10.1515/commun-2014-0029

Govani, T. & Pashley, H. (2005). Student awareness of the privacy implications when using Facebook. *(Unpublished paper presented at the "Privacy Poster Fair" at the Carnegie Mellon University School of Library and Information Science.) 9*, 1–17.

Grinnell, C. K. (2009). From consumer to prosumer to produser: Who keeps shifting my paradigm? (We do!) *Public Culture*, *21*(3), 577–598. doi:10.1215/08992363-2009-009

Gross, R. & Acquisti, A. (2005). Information revelation and privacy in online social networks. In *Proceedings of the 2005 on privacy in the electronic society* (pp. 71–80). WPES '05. New York, NY, USA: ACM. doi:10.1145/1102199.1102214

Gruber, T. (2008). Collective knowledge systems: Where the social web meets the semantic web. *Web semantics*, *6*(1), 4–13. doi:10.1016/j.websem.2007.11.011

Gutwirth, S. (2002). *Privacy and the Information Age*. Oxford: Rowman & Littlefield Publishers.

Heale, R. & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence Based Nursing*, *18*(3), 66–67. doi:10.1136/eb-2015-102129

Horváth, Z., Bogaerts, S., Sijtsema, J., & Demeyer, K. (2014). Privacy, risk, information protection and social network site-using behavior in a sample of Flemish university students. *Leuven Institute of Criminology, Tilburg University*, 1–19.

Joinson, A. N. & Paine, C. B. (2007). Self-disclosure, privacy and the Internet. In A. N. Joinson, K. Y. A. McKenna, T. Postmes, & U Reips (Eds.), *Oxford handbook of internet psychology* (pp. 235–250). Oxford University Press.

Jourard, S. M. & Lasakow, P. (1958). Some factors in self-disclosure. *The Journal of Abnormal and Social Psychology*, *56*(1), 91.

Kabacoff, R. I. (2015). *R in action: Data analysis and graphics with R*. New York: Manning.

Kaplan, A. M. & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, *53*(1), 59–68. doi:10.1016/j.bushor.2009.09.003

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference* (pp. 591–600). WWW '10. New York, NY, USA: ACM. doi:10.1145/1772690.1772751

Lewis, K., Kaufman, J., & Christakis, N. (2008). The taste for privacy: An analysis of college student privacy settings in an online social network. *Journal of Computer-Mediated Communication*, *14*(1), 79–100. doi:10.1111/j.1083-6101.2008.01432.x

Litt, E. (2013). Understanding social network site users' privacy tool use. *Computers in Human Behavior*, *29*(4), 1649–1656. doi:10.1016/j.chb.2013.01.049

Livingstone, S. (2008). Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. *New Media & Society, 10*(3), 393–411. doi:10.1177/1461444808089415

Livingstone, S. (2013). The participation paradigm in audience research. *The Communication Review, 16*(1-2), 21–30. doi:10.1080/10714421.2013.757174

Lohr, S. L. (2008). Coverage and sampling. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 97–112). New York, NY, USA: Taylor & Francis Group.

Magazine, S. L., Williams, L. J., & Williams, M. L. (1996). A confirmatory factor analysis examination of reverse coding effects in Meyer and Allen's Affective and Continuance Commitment Scales. *Educational and Psychological Measurement*, *56*(2), 241–250.

Marwick, A. E. (2012). The public domain: Social surveillance in everyday life. *Surveillance and Society*, *9*(4), 378–393.

McKeon, M. (2010). The Evolution of Privacy on Facebook: Changes in default profile settings over time. Retrieved from http://mattmckeon.com/facebook-privacy/

Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Essex: Pearson. doi:10.2307/3211488

Nigam, H. (2013). Social privacy: What does this really mean? Retrieved from http://www.huffingtonpost.com/hemanshu-nigam/facebook-privacy_b_2925334.html

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life.* Stanford: Stanford University Press.

O'Brien, D. & Torres, A. M. (2012). Social networking and online privacy: Facebook users' perceptions. *Irish Journal of Management*, *31*(2), 63–97.

O'Connor, B. T. (2014). *Statistical Text Analysis for Social Science* (Doctoral dissertation, Carnegie Mellon University).

OECD. (2007). Participative web and user-created content: Web 2.0, wikis, and social networking. Paris: *Organisation for Economic Co-operation and Development.* (2006), 74. doi:10.1787/9789264037472-en

O'Reilly, T & Battelle, J. (2009). *Web Squared: Web 2.0 Five Years On.* O'Reilly & Techweb.

Osborne, J. W. & Costello, A. B. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 1–9.

Park, M.-S., Shin, J.-K., & Ju, Y. (2015). A taxonomy of social networking site users: Social surveillance and self-surveillance perspective. *Psychology & Marketing, 32*(6), 601–610. doi:10.1002/mar.20803

Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the Horizon*, *9*(5), 1–6. doi:10.1108/10748120110424816

Raacke, J. & Bonds-Raacke, J. (2008). MySpace and Facebook: Applying the uses and gratifications theory to exploring friend-networking sites. *Cyberpsychology & Behavior: The impact of the Internet, multimedia and virtual reality on behavior and society*, *11*(2), 169–74. doi:10.1089/cpb.2007.0056

Raubenheimer, J. (2004). An item selection procedure to maximise scale reliability and validity. *SA Journal of Industrial Psychology; Vol 30, No 4 (2004)*. Retrieved from http://www.sajip.co.za/index.php/sajip/article/view/168

Regan, P. M. (1995). *Legislating privacy: Technology, social values, and public policy.* Chapel Hill & London: The University of North Carolina Press.

Riffe, D., Lacy, S., & Fico, F. (2014). *Analyzing media messages: Using quantitative content analysis in research* (3rd ed.). New York: Routledge.

Ritzer, G., Dean, P., & Jurgenson, N. (2012). The coming of age of the prosumer. *American Behavioral Scientist*, *56*(4), 379–398. doi:10.1177/0002764211429368

Roberts, P. (2006). Reliability and validity in research. *Nursing Standard*, *20*(44), 41–45. doi:10.7748/ns2006.07.20.44.41.c6560

Shao, G. (2009). Understanding the appeal of user-generated media: A uses and gratification perspective. *Internet Research*, *19*(1), 7–25. doi:10.1108/10662240910927795

Sloan, R. H. & Warner, R. (2013). Big data and the 'new' privacy tradeoff. *Kent College of Law Research Paper*, *33*. doi:10.2139/ssrn.2306071

Solove, D. J. (2002). Conceptualizing privacy. *California Law Review*, *90*(4), 1087–1155. doi:10.1145/1929609.1929610

Spence, I. (2005). No humble pie: The origins and usage of a statistical chart. *Journal of Educational and Behavioral Statistics*, *30*(4), 353–368.

Tokunaga, R. S. (2011). Social networking site or social surveillance site? Understanding the use of interpersonal electronic surveillance in romantic relationships. *Computers in Human Behavior*, *27*(2), 705–713. doi:10.1016/j.chb.2010.08.014

Tokunaga, R. S. (2016). Interpersonal surveillance over social network sites. *Journal of Social and Personal Relationships*, *33*(2), 171–190. doi:10.1177/0265407514568749

Trottier, D. (2010). *Mutual Augmentation of Surveillance Practices on Social Media* (Doctoral dissertation, Queen's University, Canada).

Trottier, D. (2012a). Interpersonal surveillance on social media. *Canadian Journal of Communication*, *37*, 319–332.

Trottier, D. (2012b). *Social media as surveillance: Rethinking visibility in a converging world.* Surrey: Ashgate.

Trottier, D. (2015). Interpersonal surveillance and user-led policing on social platforms. *The Social Life of Big Data Symposium.* Retrieved from http://ro.ecu.edu.au/slbd/2/

Tufekci, Z. (2008). Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology & Society*, *28*(1), 20–36. doi:10.1177/0270467607311484

Tuunainen, V., Pitkänen, O, & Hovi, M. (2009). Users' awareness of privacy on online social networking sites - Case Facebook. *BLED 2009 Proceedings*, *42*.

van Dijck, J. (2009). Users like you? Theorizing agency in user-generated content. *Media, Culture & Society*, *31*(1), 41–58. doi:10.1177/0163443708098245

Warren, S. D. & Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review*, *4*(5), 193–220. doi:10.2307/1321160

Waters, S. & Ackerman, J. (2011). Exploring privacy management on Facebook: Motivations and perceived consequences of voluntary disclosure. *Journal of Computer-Mediated Communication*, *17*(1), 101–115. doi:10.1111/j.1083-6101.2011.01559.x

Whiting, A. & Williams, D. (2008). Why people use social media: A uses and gratifications approach. *An International Journal Aslib Proceedings Georgios Tsimonis Sergios Dimitriadis Marketing Intelligence &amp; Planning Journal of Research in Interactive Marketing*, *16*(4), 362–369. doi:10.1108/QMR-06-2013-0041

Williams, L. D., Crittenden, L. V., Keo, T., & McCarty, P. (2012). The use of social media: An exploratory study of usage among digital natives. *J. Public Affairs*, *12*, 127–136. doi:10.1002/pa.1414

Wisniewski, P. J., Knijnenburg, B. P., & Richter, H. R. (2017). Making privacy personal: Profiling social network users to inform privacy education and nudging. *Journal of Human Computer Studies*, *98*(September 2016), 95–108. doi:10.1016/j.ijhcs.2016.09.006

Wyrwoll, C. (2014). *Social Media Fundamentals, Models, and Ranking of User-Generated Content*. Hamburg, Germany: Springer Vieweg. doi:10.1007/978-3-658-06984-1

Young, A. & Quan-Haase, A. (2013). Privacy protection strategies on Facebook. *Information, Communication and*, *4462*(April). doi:10.1080/1369118X.2013.777757

Zwitter, A. (2014). Big data ethics. *Big Data & Society*, *1*(2), 2053951714559253. doi:10.1177/2053951714559253

Appendix A

Dear participant,

This survey research is about the level of social privacy awareness on Facebook among UNIVERSITY STUDENTS IN THE NETHERLANDS. In brief, social privacy is about controlling access to personal information data and attempts to preserve privacy by keeping others' unwanted audience away.

The survey will not be longer than 10 minutes and all your responses are completely anonymous. You can only take the survey once. Please read the survey questions carefully and provide an answer to every question. Questions marked with an asterisk (*) are required. If you have any questions about the survey, please email me: 455219my@student.eur.nl

The survey will be closed on 1st June, 2017.

Your answers will contribute to my academic research and make graduate degree possible. Thank you so much for your participation.

Metin Yazici
MA Media, Culture & Society

Q2.1 What is your gender?*

○ Male (1)
○ Female (2)
○ Other (3)

Q2.2 What is your age?*

____

Q2.3 What is your nationality?

____

Q2.4 What do you study?*

____

Q2.5 What is your level of education?*

○ 1st year (1)
○ 2nd year (2)
○ 3rd year (3)
○ Masters (4)
○ Other (5)

Q3.1 Do you have a Facebook profile?*
○ Yes (1)
○ No (2)

Q3.2 Have you ever had a Facebook profile?*
○ Yes (1)
○ No (2)

Q4.1 How long have you been on Facebook? Since...*
○ 2004 (1)
○ 2005 (2)
○ 2006 (3)
○ 2007 (4)
○ 2008 (5)
○ 2009 (6)
○ 2010 (7)
○ 2011 (8)
○ 2012 (9)
○ 2013 (10)
○ 2014 (11)
○ 2015 (12)
○ 2016 (13)
○ 2017 (14)

Q4.2 How many hours do you spend on checking your newsfeed on Facebook per day (average)?*
○ 0-1 (1)
○ 1-3 (2)
○ 4-8 (3)
○ 9-12 (4)
○ 12+ (5)

Q4.3 In general, how often do you go on Facebook?*
○ More than twice a day (1)
○ Once a day (2)
○ Twice a week or more (3)
○ Once a week (4)
○ Once a month (5)
○ Never (6)

Q4.4 On average, how often do you share content on Facebook? (For instance update status, add photo, check-in somewhere etc.)*

○ More than twice a day (1)
○ Once a day (2)
○ Twice a week or more (3)
○ Once a week (4)
○ Once a month (5)
○ Never (6)

Q5.1 Friends (with capital F) on Facebook refers to your connected network on Facebook including all the users you interact with, which is different than everyday understanding of friends.

Q5.2 How many Friends do you approximately have on Facebook?*

○ 1-100 (1)
○ 101-200 (2)
○ 201-300 (3)
○ 301-500 (4)
○ 500+ (5)

Q5.3 What would you consider as the level of connection you have with your Friends on Facebook? Please think and answer by the term of degrees.*

	1 (1)	2 (2)	3 (3)	4 (4)
Close friends:* (1)	○	○	○	○
Acquaintances:* (2)	○	○	○	○
Distant friends:* (3)	○	○	○	○
People only met on Facebook:* (4)	○	○	○	○

Q6.1 What kind of personal information do you have shared in your Facebook profile, and how complete and accurate is it?*

	I share this information complete and accurate (1)	I share this information but it is not complete or accurate (2)	I don't share this information (3)
Full name (1)	○	○	○
Date of Birth (2)	○	○	○
Hometown or City (3)	○	○	○
E-mail address (4)	○	○	○
Telephone Number (5)	○	○	○
Home address (6)	○	○	○
Relationship status (7)	○	○	○
Biography (8)	○	○	○
Family members (9)	○	○	○
School or employment (10)	○	○	○
Political views (11)	○	○	○
Religion (or content related to it) (12)	○	○	○
Sexual orientation (13)	○	○	○
Partner's name (14)	○	○	○
Family's name (15)	○	○	○
Photos of you (16)	○	○	○
Photos of your friends (17)	○	○	○
You travelling status (for example, going on vacation) (18)	○	○	○
Opinions about your job, school, family (19)	○	○	○
Places you visit (Check-in location) (20)	○	○	○
Favorite music, book, movie etc. (21)	○	○	○
Important Life Events (22)	○	○	○

Q7.1 To what extent do you agree the following on your Facebook account?*

	Strongly disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Strongly agree (5)
I worry about people using Facebook to try to discover more information about me (1)	○	○	○	○	○
I often think about who might be reading what I post and yet not responding (2)	○	○	○	○	○
I do not give much thought to whether people are actively monitoring what I post (3)	○	○	○	○	○
I think about the extent to which people may be creeping on my Facebook page (4)	○	○	○	○	○
I worry about who may be engaging in prolonged scrutiny of my Facebook page (5)	○	○	○	○	○
I often think about who might be reading my Facebook content and want to go undetected (6)	○	○	○	○	○
I often scrutinize what information I post on Facebook (7)	○	○	○	○	○
I think about how comfortable I am with the level of exposure my Facebook content might bring (8)	○	○	○	○	○
I do not worry about people trying to use Facebook to creep on me (9)	○	○	○	○	○
I do not think about who may be constantly monitoring my Facebook page (10)	○	○	○	○	○

Q8.1 Privacy settings allows user to control who can see the information. The default recommended, privacy settings of Facebook is set to Public, which makes your content open to all users in the network.

Q8.2 How often do you perform the following things on your Facebook account?*

	Never (1)	Sometimes (2)	About half the time (3)	Most of the time (4)	Always (5)
I send private e-mail messages instead of posting to a Friend's wall to restrict others from reading the message (1)	❍	❍	❍	❍	❍
I usually go offline on Facebook chat (2)	❍	❍	❍	❍	❍
I exclude personal information on Facebook to restrict people I don't know from gaining information about me (3)	❍	❍	❍	❍	❍
I untag myself from images and/or videos posted by my contacts (4)	❍	❍	❍	❍	❍
I ask my Friends to remove tags from my posts or photos (5)	❍	❍	❍	❍	❍
I delete posts from others to my Facebook wall to restrict others from viewing/reading the post (6)	❍	❍	❍	❍	❍
Certain contacts on my Facebook site only have access to my limited profile (7)	❍	❍	❍	❍	❍

	Department		
1	International law	89	BA in Psychology
2	Law	90	Bedrijfseconomie
3	Media, culture and society	91	Business Economics
4	MA Media Studies - Media Culture and Society	92	International Business Administration
5	Media	93	IBA
6	Media & communication	94	BSc International Business Administration
7	Educational sciences	95	Business
8	French studies	96	Law
9	Architecture	97	MSc Business Information Management (BIM)
10	International Bachelor of Economics and Business Economics	98	LLM in MARITIME AND TRANSPORT LAW
11	Arts and Culture	99	Econometrics and Operational Research
12	Media, culture and society	100	strategic manament
13	Communication & Media	101	IBA
14	Media Technology	102	Organizational change and consulting
15	Psychology	103	Graphic Design
16	Neuroscience	104	ELECTRICAL ENGINEERING
17	Media Communication	105	IBA
18	Economics	106	MASTER MARITIME AND TRANSPORT LAW
19	Stochastics and mathematical finance	107	Law
20	International arts and culture studies	108	Pre master media culture and society
21	Marketing	109	Leisure management
22	International Relations & Diplomacy	110	Bedrijfskunde MER
23	Media and Communication	111	facility management
24	International marketing	112	Philosophy
25	Hbo	113	International Business Administration
26	Commerciele economie	114	Public Administration
27	Commerciële economie	115	Liberal Arts and Sciences
28	Politics	116	Law
29	International law	117	Behavioural Economics
30	Commerciële Economie	118	IBA
31	Master Media & Business	119	Liberal Arts and Sciences
32	Software development	120	Global management of international social challenges
33	Computer Engineering	121	International Bachelor Psychology
34	Sociology	122	Social science, leisure studies
35	Arts and Culture studies	123	Bestuurskunde/public administration
36	Master in international economics	124	Economics and Business
37	Classic Music, Voice	125	Econometrics
38	International Bachelor Arts and Culture Studies	126	Bedrijfskunde
39	commercial and company law	127	Commercial and Company Law
40	International Communication & Media	128	Computer Science
41	Accountancy	129	Business information management
42	Law	130	Applied Physics
43	public relations	131	Aerospace engineering
44	Law	132	Trade Management focused on Asia
45	Law	133	Finance
46	Psychology	134	International Business Administration
47	Latin American Studies	135	Embedded Systems
48	journalisme	136	Sociology
49	Medical University of Warsaw; Science of public health	137	Technology and Operations Management
50	Critical and Cultural Studies	138	IBA
51	Business and Management	139	Health care management
52	Media, Culture, & Society	140	Finance
53	IBA	141	International Business Administration
54	Biology	142	IBA
55	IT	143	International Communication and Media
56	IBCoM	144	Marketing Management
57	Industrial Engineering	145	Business Information Management
58	Math	146	Media business
59	Rotterdam	147	Finance & Investments
60	Econometrics and Management Science	148	Finance
61	Law	149	IBEB
62	law	150	Economics & Business Economics
63	Law	151	MA Cultural Economics and Entrepreneurship
64	International Business Administration	152	Strategic Management
65	Business administration	153	Ibcom
66	Business	154	Business administration
67	IBA	155	International Business Administration
68	Business administration	156	medicine
69	Medicine	157	Psychology
70	Chemical engineering	158	IBA
71	Chemical engineering	159	Finance
72	IBCoM	160	BA
73	International Psychology	161	MSc Economics and Business
74	Psychology	162	IBEB
75	Sociology	163	Econometrics and Operations Research
76	Bussines Adimistiration	164	Economics
77	international economics	165	Economics and Business Economics
78	Bsc in Public Administration (Management of International Social Challenges)	166	Business Information Management
79	Economics	167	International Bachelors of Economics and Business
80	Liberal Arts & Sciences	168	Economics and Business Economics
81	Commercial and Company law	169	IBA
82	Marketing	170	Business
83	IBA	171	Economics
84	Commercial and company law	172	chemistry
85	Law	173	Law
86	Law	174	Physics
87	Law	175	Physics
88	Commercial Law	176	Aerospace

The list of study departments of respondents.

Appendix B

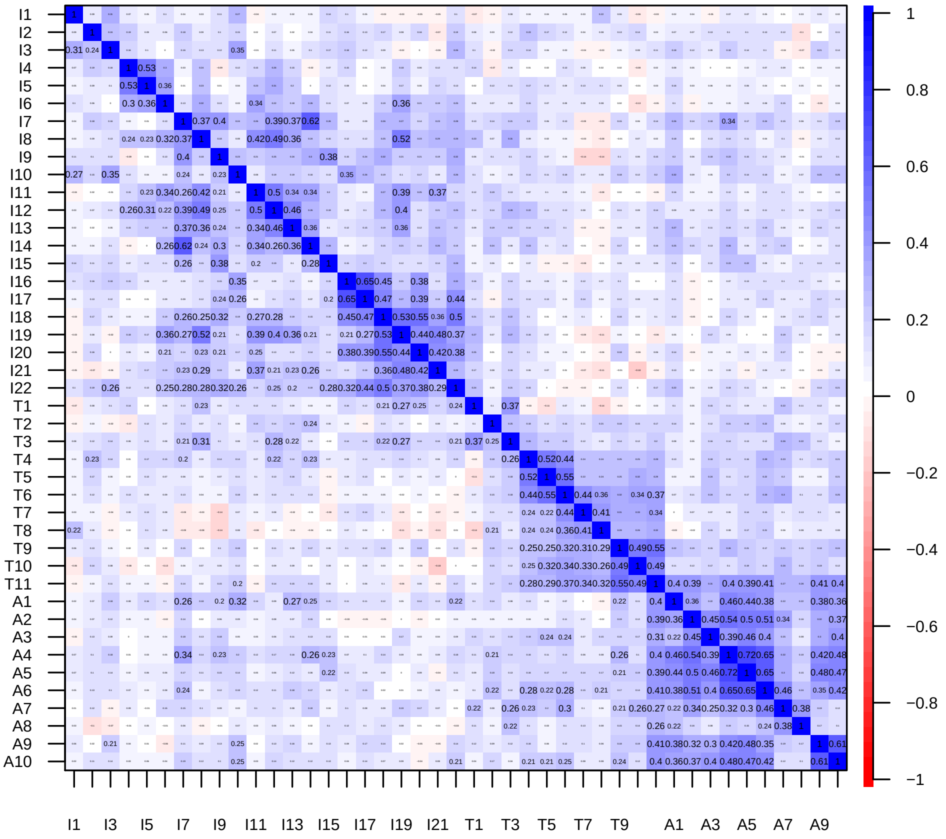


Figure 1B: Correlogram of validity for correlation levels and factor matrices

Construct	Cronbach's α
Basic Information Disclosure	.775
Appearance Information Disclosure	.814
Technological Privacy Tools	.819
Awareness of Social Privacy	.875

Table 1B: Results of Cronbach's alpha (α) of the constructs

	Rotated Components			
Item	1	2	3	4
I7	.239		.704	
I9	.191		.460	.277
I11			.656	.108
I12		.137	.678	.118
I13	.116		.636	
I14	.145		.690	
I16				.759
I17		.113		.798
I18			.276	.764
I19			.535	.468
I20			.205	.676
I22	0.153		.263	.635
T4		.588	.280	
T5		.665	.173	
T6	.103	.753	.136	
T7		.663	-.155	
T8		.618	-.205	
T9	.233	.576		.150
T10	.139	.630		
T11	.529	.548		
A1	.617		.199	.124
A2	.718			-.131
A3	.563	.182	.125	
A4	.806		.187	
A5	.815	.126		
A6	.720	.215	.102	
A9	.662			.130
A10	.681	.159		
Eigenvalues	4.493	3.395	3.264	3.101
Proportion S^2	.160	.121	.117	.111
Cumulative S^2	.160	.282	.398	.509

Table 2B: Rotated factor loadings based upon correlation matrix