THE ABILITY OF USERS OF SOCIAL NETWORKING SITES TO ACCURATELY PERFORM THE PRIVACY CALCULUS

An empirical study on the disclosure abilities of Facebook users

Master thesis report Business Administration
Information Management
School of Management and Governance
University of Twente, Enschede

Stefan van de Worp (S1242156)
17-11-2014

Supervisors:
Dr. A.B.J.M Wijnhoven
Dr. L.C.P Broos LLM

UNIVERSITY OF TWENTE.

ABSTRACT

This report discusses the results of an empirical study on the competences, skills and knowledge of users of Facebook in personal information disclosure decisions. A survey measured the competences that determine the ability to accurately perform the privacy calculus. The competences for the disclosure decision consist of the ability to estimate perceived benefits, perceived privacy risks and perceived trust on Facebook. The tested skills consist of privacy policy and privacy control skills. The knowledge is determined by the awareness of data exploitation.

The study is performed with Dutch adult Facebook users as the unit of observation. In the research the competence levels of the users are determined by comparing their survey results with a golden standard. The study provides a transparent scale by with the users are categorized in competence, skills and knowledge levels. Linear regression analysis is used to explore the field of variables that determine the competence levels and skill levels to perform the privacy calculus.

The results of the study indicate that Dutch adult Facebook users are categorized a sufficiently competent in estimating the benefits and risks of data disclosure situations on Facebook. However, they lack the specific skills and knowledge which are required to use the information service without endangering the privacy of themselves and others.

Keywords: privacy, social media, Facebook, competences, privacy calculus

TABLE OF CONTENTS

1 INTRODUCTION

An increasing number of online social networking services enter the information market to battle for the attention of internet users. Internet users have multiple possibilities to interact, socialize and play games on the internet. A broad variety of those services offer endless functionalities to attract large groups of customers. The networks persuade the internet users with free benefits and enjoyment to join the services. The increased sharing of personal information on social networks bears benefits as well as privacy risks. The services are eager to gather more and diverse information about their users, while the users' concerns with privacy increase (Malhotra, 2004b).

The mental cost / benefit consideration between the benefits and privacy risks is called the privacy calculus (Dinev & Hart, 2006). Social networking sites (SNS) ask users to disclose personal information. But to what extent are the users of SNS capable of rationally deciding to disclose personal information? The main research question of this research is: **To which extent are Dutch adult Facebook users able to accurately perform the privacy calculus?**

The privacy calculus (PC) gives insights in the antecedents and dimensions of the willingness to disclose information on the internet. In figure 1.1 the basic conceptual model of this research is illustrated. This model is based on the PC model of Krasnova & Veltri (2010).

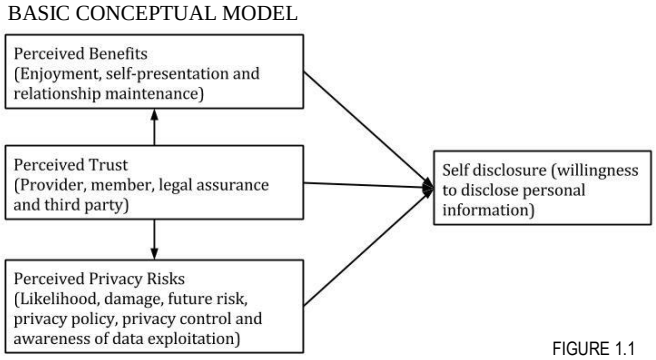

FIGURE 1.1

This research measures to which extent the users have the ability to accurately estimate the dimensions and variables that are required to accurately perform the privacy calculus.
Users of Facebook may be categorized in competence levels of performing the privacy calculus. This research presents the competence levels of users in performing the estimation of those antecedents.

The scientific relevance of this study creates insights in the ability to perform the privacy calculus for the users of social networks. This empowers users to make well-considered disclosure decisions, since they are exposed to privacy risks more often. Currently the society is at a crossing to shape the future of privacy-related information disclosure decisions. The growth of the internet remains to gain speed[1]. Big data and 'The internet of Things' (IOT) thrive this growth by creating endless possibilities with the data. IOT improves the connectivity of devices and is currently embodied by

[1] http://www.statista.com/statistics/267181/forecast-of-consumer-internet-traffic-through-e-mail-and-web-usage/).

wearables like Smart Watches, Google Glass, medical devices and also Smart Thermostats and smart cars. Big data drives predictive analysis to support these instruments. The growth of these phenomena has a flipside. The privacy concerns of individuals grow with the introduction of these smart technologies. Predictive analysis (by data-mining) can harm privacy due to data inaccuracy (Che, Safran & Peng, 2013), out-of-context-analysis, apophenia, errors from combining data sets (Boyd & Crawford, 2012) and the filter bubble (Parsiser, 2011).
From a business perspective the increased use of Xquick and DuckDuckGo illustrate the concerns regarding privacy among search engines, the growth of browser add-ons Ghostery and DoNotTrackMe illustrate the desire to browse without being tracked and the founding of Ello illustrates that the design of SNS can be performed with privacy taken in account.
The EU is currently in the process of reforming the Data protection legislation introduced in 1995. 'Current rules need to be modernized – they were introduced when the Internet was still in its infancy'. 'Rapid technological development and globalization have brought new challenges for data protection' ('Why do we need an EU data protection reform?', 2014). New legislation is in the process of being approved by the European Commission. The key changes in the modernized legislation are: a right to be forgotten, explicit consent before data processing, right of data portability, privacy by default, privacy by design and clear and transparent data operations. This single set of rules stimulates privacy protection and has economic advantages. Old and new legislation already impact business in a significant way. The EU Court of Justice enacted the right to be forgotten on the 13th of May 2014, which lead to the introduction of the 'Right to be forgotten' form by Google. Facebook already implemented privacy by default features and simplified the policy as a reaction to the upcoming modifications in legislation.
The above mentioned trends and developments illustrate the crossing the society is at.
It becomes clear that privacy is a hot topic in science, legislation-debates and modern technology. The results of this research aim to support to shape the future regarding privacy-related information disclosure.

The conducted research mainly consists of two parts. First the conceptual model of the privacy calculus is corroborated. Secondly, the estimation competences, skills and knowledge of the users are categorized.

2 THEORETICAL FRAMEWORK

Online information services are highly integrated in the daily lives of people. We search for information on Google, buy products at Amazon and socialize with Facebook and Twitter. More and more companies collect more data of their users which they try to commercialize. With the ongoing growth of online information services a double cutting sword principle emerges for the users. On the one hand disclosing personal information benefits the users by more suiting services, while on the other hand their privacy is endangered. Xu et all (2011). call this the profiling – privacy paradox. This paradox also befalls to governments. A government strives to protect their citizens from harmful or unsafe events by monitoring their activities; on the other hand a government aims to respect the call for privacy by the same citizens. This Orwellian discussion balances between these extremes.

Dinev and Hart (2006) put this contradiction in a bigger picture known as the privacy calculus. Prior to personal information disclosure users go through a decision process. The theory of planned behavior (TPB) is an influential theory that predicts the decision behavior of individuals. The privacy calculus gives insights in antecedents for these decisions.

2.1 Privacy Calculus

The privacy calculus is the mental trade-off that precedes personal information disclosure decisions. In the cognitive process before this decision a rational cost / benefit trade-off is performed. The privacy calculus was firstly coined by Laufer & Wolfe in 1977. At that time the term calculus of behavior was used more frequently. Laufer & Wolfe created a multidimensional developmental view on privacy. They state that privacy is a situational dependent social issue. It is described by the self-ego, the environment and the interpersonal relationships.

Laufer & Wolfe (1977) state that the cognitive trade-off for individuals consists of the perceived benefits and the estimated risks of their disclosure decision. It is labeled as a ‘calculus’ because the individual weights the prior named variables before disclosing. An individual will disclose personal information when he perceives that the overall benefits of disclosure are at least balanced by, if not greater than, the assessed risk of disclosure (Culnan & Bies, 2003). Dinev & Hart (2006) and Culnan and Armstrong (1999) have elaborated on the antecedents of the privacy calculus. Culnan and Armstrong (1999) further introduce the influencing factor of trust. These researchers focus on the disclosure of personal information the internet.

Privacy risk

Privacy in the context of the privacy calculus often refers to privacy risks. Risks are hard to quantify since they are a probability of the occurrence of an event. Accurately determining the weight or likelihood of a risk is complex, if not impossible. Privacy risks have social, economic, legislative and informational perspectives.

The basic perception of privacy is known as the ability to seclude the individual or information about the individual from others. People wear clothes and build fences around their houses to satisfy this type of privacy. Having the choice of exposure is a great deal within this concept. This interpretation of privacy is part of the Universal Declaration of Human Rights (1950) and states: *'respect for private and family life. Everyone has the right to respect for his or her private and family life, home and correspondence.'*
This basic right to privacy became more specific with the rise of the internet. The EU created legislation that protects privacy on a more economical and informational manner. Legislation on privacy is categorized by the EU within the protection of personal data. Personal data is according to the Council of Europe '*any information relating to an identified or identifiable individual (data subject)'*. The privacy of individuals thus switched to a focus of informational privacy. Current campaigns (Aangenomen Teksten EP 12-3-2014) regarding data protection provide examples: *'Any information relating to an individual, whether it relates to his or her private, professional or public life. It can be anything from a name, a photo, an email address, your bank details, your posts on social networking websites, your medical information, or your computer's IP address.'*

According Yates and Stone (1992) privacy risk is 'the possibility of loss' which is 'an inherently subjective construct (Chiles and McMackin, 1996).' Privacy risk also includes the misuse of personal information, such as insider disclosure or unauthorized access and theft (O'Brien 2000, Rindfleisch 1997). Featherman and Pavlou (2003) specify by stating that privacy risk is 'potential loss of control over personal information, such as when information about you is used without your knowledge or permission.' The likelihood is estimated by the users of SNS as part of the privacy calculus. The estimated damage (impact) is together with the likelihood the determining factor of a privacy risks (Krasnova, Kolesnikova & Gunether, 2009).

Perceived benefits
Rosen and Sherman (2006) state that enjoyment is central to the use of SNS, while Boyd (2007) argues that self-presentation is the main driver of the use of SNS. The third reason to disclose information is labeled as relationship maintenance. Different terms like improved connectedness (Hogben, 2007), self enhancement, emotional support (Koroleva and all., 2011) and networking value (Koroleva and all., 2011) are nearly similar as the previously mentioned factors but just differently labeled. A study by Sheldon (2008) explores more motives for joining Facebook; most factors are covered by the earlier mentioned terms. However additional motives for participating on Facebook are 'to pass the time, for entertainment, sense of virtual community, coolness and/or companionship'. The privacy calculus is an information privacy activity to interact between individuals, groups and the provider.
Social capital is also often mentioned when discussing the benefits of social interaction. In 1988 Coleman refers to social capital as 'the resources accumulated through the relationship between people'. At that moment in time the social capital term was not developed for online interactions and friendships, although it does apply to it. More recent definitions of social capital are closer related to online networks. Social capital is defined by Bourdieu and Wacquant (1992) as 'the sum of the resources, actual or virtual, that accrue to an individual or a group by virtue of possessing a durable network of more

or less institutionalized relationships of mutual acquaintance and recognition'. The online resources, for instance those from Facebook, are embodied by likes, comments and tags.

Perceived Trust

Perceived trust is the third variable which influences the willingness to disclose. Dinev and Hart (2005) see the trust in the internet as of great importance to the willingness to disclose. Internet trust is 'trust beliefs reflecting confidence that personal information submitted to enticement beliefs websites will be handled competently, reliably, and safely' (Dinev & Hart, 2005). They conclude that Internet Trust is indirectly related to a willingness to disclose personal information and that perceived internet trust is directly related to perceived privacy risks. Internet trust is therefore indirectly incorporated in the privacy calculus model.

The online organization actively needs to develop and maintain trust relationships with their customers. There is a great difference between trust in the internet and trust in organizations. Kehr, Wentzel and Mayer speak of institutional trust (trust in organizations), which refers to an individual's confidence that the data-requesting stakeholder, or medium will not misuse his or her data (Andersons and Agarwal, 2011). It has been found to be related to privacy concerns (Malhotra et al, 2004), and intentions to disclose information. The definition by Andersons and Agarwal from 2011 specifies the concept: 'Trust in the data-collection electronic medium as a pre-existing cognitive factor that may be affected by situational variables such as beliefs about the stakeholder requesting the information.'

According to Smith et al (2011) the relationship between information privacy and trust has not been modeled consistently in the literature. Some authors treat trust in the relationship as an antecedent and others as the result of a privacy concern.

Malhotra, Kim and Agarwal (2004) treat trusting beliefs as a result of privacy concerns. They proposed the Internet Users' Information Privacy Concerns model (IUIPC). In this model the privacy concerns of internet users are determined by *collection*, *control* and *awareness* which the users perceive. The *collection* of personal information is fundamental to privacy concerns. Collection therefore is 'the degree to which a person is concerned about the amount of individual-specific data possessed by others relative to the value of benefits received.' The collection of data by Facebook is stated in the privacy policy. The ability to estimate the meaning of the statements of the privacy policy indicates the skill of the collection.
The meaning of the concept *control* regards to the possibility to approve, modify or exit (i.e. opt-out) the service (Caudill and Murphy, 2000). These options are embodied by control settings. The privacy settings of Facebook represent the concept of control. Awareness is the level of concerns regarding the awareness of organizational information privacy practices (Culnan, 1995). The same study also found evidence that the type of information that is requested by the organization also significantly influences the behavioral intention to disclose personal information. The managerial implications that result from the study of Malhotra et all. (2004a) provide two strategies for SNS

providers: empower users with control or increase awareness in the data handling operations.

Krasnova & Veltri (2010) narrow the concept of trust regarding information systems down to three moderating factors: provider trust, member trust and trust in legal assurance. They state that trust in the provider and in the members of the SNS are the largest contributors to the influence trust on the willingness to disclose. The trust in legal assurance is also a heavy weighting variable. Trust in SNS members is defined as the non-opportunistic behavior, benevolence and trustworthiness of other persons from the social group. Trust in legal assurance surpass the provider / member domain and is defined as the confidence in the legal framework that protects the members (e.g. governmental legislation).

The partners of Facebook (third parties) are also data-handling entities. In this research the *competence in estimating the trust in third parties* is an additional competence regarding trust. The SNS provider is able to increase the trust by implementing operational fairness. Fair operational practices are procedural practices that increase the control over the personal information (Culnan & Armstrong, 1999).

Krasnova and Veltri (2010) stress the fact that cultural aspects are of major influence on the perception of the trustee. They found that the influence of the trust in the provider and the trust in the members is more present in the USA than in Germany. Krasnova and Veltri found a positive relationship between the uncertainty avoidance and the privacy concerns of the individual. The individualism is positively correlated with the enjoyment and the trust factors (provider and members). As established by Hofstede[2] (2001), the cultural characteristics influence the behavior of individuals significantly. The disclosure behavior for instance is greatly influenced by the level of individualism, masculinity and uncertainty avoidance. In figure 2.1 the cultural characteristics of The Netherlands, China and the United States are illustrated.

CULTURAL CHARACTERISTICS INFLUENCING DISCLOSURE

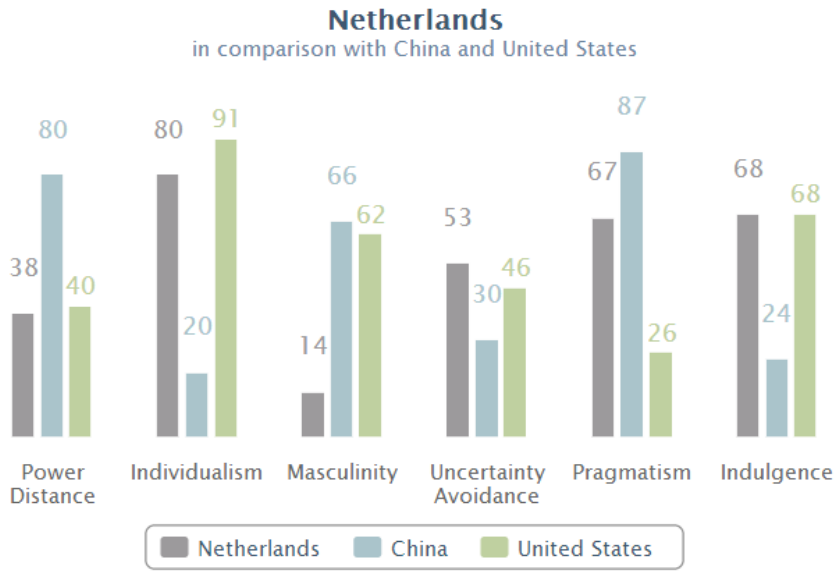

FIGURE 2.1

[2] http://geert-hofstede.com/netherlands.html

The levels of masculinity, individualism and uncertainty avoidance deviate between the countries. This indicates different cultural characteristics, thus the results of this research cannot be generalized to the entire Facebook population.

It is clear that trust can indirectly positively influence the willingness to disclose personal information. Gefen et all (2003) add that trust is a multidimensional construct and it has been defined in numerous ways. Dinev & Hart (2006) define trust from an e-commerce perspective: 'a set of specific beliefs about another party that positively influences an individual's intention to conduct an online transaction. These beliefs embody the expectation that another party will not engage in opportunistic behavior.' In transactions and operations domain Jarvenpaa (2000) explains trust as a set of beliefs including the expectations that an online vendor would keep the best interest of the consumer and its promises to them in mind.

The collection of information

Intentions to disclose or intentions of self-disclosure are the known alternative descriptions for the willingness to disclose. There is a difference between actually disclosing and measuring the attitude towards disclosing. The willingness to disclose is a matter of attitude, while disclosure of information concerns behavior.

When disclosing personal information a distinction is made between explicit and implicit data (Gauch, Speretta, Chandramouli, Micarelli 2007). The collection of personal information from users of SNS is executed by two prevailing methods: implicit and explicit data gathering. Explicit data collection is executed through the direct intervention of the concerning users. Implicit data collection is the data collection performed by software agents that monitor the user activity (Gauch, Speretta, Chandramouli, Micarelli 2007).

The tool often used for explicit data gathering is a HTML-form or a similar instrument. To create a profile the user is asked to enter personal information about himself like: name, place of birth and hobbies. The user can add information to his profile and has the ability to change it, which makes it dynamic. Collecting explicit data about the user has several advantages and disadvantages. The advantages for the user are that only information is gathered which the user explicitly disclosed. This positively influences control and awareness. On the flipside this is a disadvantage for the platform provider. The platform provider explicitly asked for information and users have the possibility not to disclose it or disclose incorrect information. The increased control entails an increased burden for the users. Extra actions of the users are required when constructing a profile.

Implicit information is gathered in a different manner. Software agents (e.g. apps, browsers, desktop software, OS) collect information about the user by analyzing the online behavior of the individual. The users are not directly informed about this type of information collection. A browser or OS provides access to search logs, browser cache and browsing history. The limitation for providers to explicitly ask for personal data is replaced by agents that constantly provide new data. The implicit gathered data supports to construct a more complete and correct profile of the individual, which is not always visible for the user (Che, Safran, & Peng (2013). The user is relieved of the burden to disclose information, however on the other hand loses control of the gathered data.

The profiles that companies constructed of their users also are more dynamic by the continuous gathering of personal data and therefore have higher validity. Explicit data gathering bears the disadvantages that the collected data often only constructs a static profile, while the preferences and behavior of the user changes over time. Implicit and explicit data gathering methods are simultaneously deployed 'an even more accurate and dynamic profile can be constructed with the combination of methods' (Gauch, Speretta, Chandramouli, Micarelli 2007).

A second distinction is made between self-disclosed account information and information by usage. Within Facebook a fine line exists between account information and information by usages. Account information is basic information that is explicitly stated on the profile page of the individual. Information by usage is additional collected information that is not needed to fundamentally identify the individual and is gathered by using core features of the service by the user or others.

Dinev and Hart (2006) do not differentiate between types of information. They define personal information in the context of 'information required to complete transactions on the internet.' The study of Dinev and Hart (2006) is mainly written from an economical perspective. Although transactions in an e-commerce setting and collaborating in social media is not exactly equal, the definition is be used for personal information disclosure on the internet. Krasnova and Veltri (2010) argue that self-disclosure is the amount of information shared on a user profile as well as in the process of communication with others.

2.2 Limitations to rationality of the privacy calculus

Decision theories and the privacy calculus originate from the theory of reasoned action (TRA) and the additional theory of planned behavior (TPB). The study by Ajzen and Fishbein (1980) resulted in the TRA which states a framework that explains behavioral intention and performance. Theories that explain predicting behavior and explain the process of decision making are of great importance to the privacy calculus since the calculus explains disclosure behavior and the process of decision making.

TTHEORY OF PLANNED BEHAVIOR MODEL

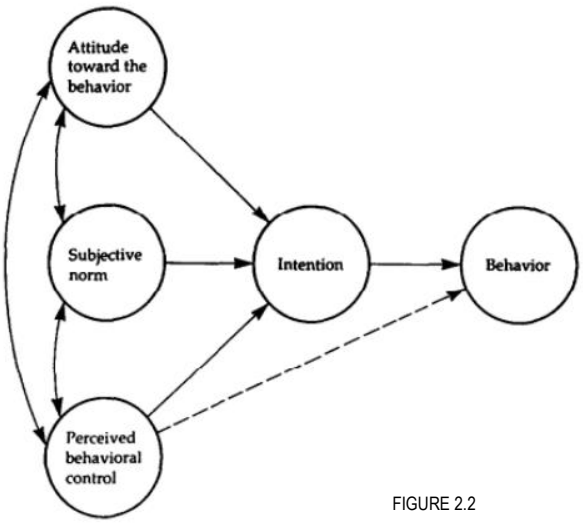

FIGURE 2.2

The goal of TPB is doing the complex task of explaining and predicting human behavior. The TPB strengthens the predictive power of human behavior. Predicting the disclosure decision of users of SNS is valuable those services. Although, it is unclear if the decision

to disclose personal information can be predicted. The theory of planned behavior suggests that the intended behavior can be predicted while the bounded rationality theory argues lack of rationality to do so. Users are assumed to act on the basis of rationality while evidence of intuition is also present.
Ajzen states that the intention is based on the attitude towards a behavior, the subjective norm and the perceived behavioral control (Figure 2.2).
The attitude towards the behavior is defined by Ajzen (2011) as 'the degree to which a person has a favorable or unfavorable evaluation or appraisal of the behavior in question.' The subjective norm refers to the perceived social pressure to perform or not to perform the behavior. The perceived behavioral control is the level of perceived complexity of performing the behavior and integrates past experiences as well as expected obstacles. When a sequence of positive past experiences is present the decision to disclose becomes more a decision based on intuition. Kahneman & Tversky (1974) state that users make disclosure decisions on an automatic-pilot. Past experiences play a larger role in frequent disclosure. The awareness per disclosure therefore differs.

A limitation of the TPB is that the theory hardly entails the influence of emotional states. Individuals in stressful situations (e.g. fear or anger) make different decisions than individuals who are not in these situations. The rational intention can be different from the behavior that is performed. Even the slightest mood changes (positive or negative) are not incorporated in the model (Ajzen, 2011). This fact co-birthed the theory of bounded rationality.

The theory of bounded rationality (TBR) by Herbert Simon provides a counterbalance to the earlier mentioned TPB. The TBR states that individuals are not entirely capable to make the best decisions. The TBR states that individuals are not entirely capable to make rational decisions since they are not able to assimilate and digest all the information that is needed to perform a certain action (Simon, 1955). Besides that the individuals often are not able to access the complete set of information, they are also restricted by their cognitive limits. The time between an action and a consequence is often long or unclear, which limits rationality.
The information scarcity results in limited decisions making. Subsequently, an abundance or complete set of information results in the inability to digest all information because of the limits in attention from the individual. The cognitive limit refers to the inability of individuals to deal with the information and knowledge at hand.
The element of time provides the third argument for bounded rationality. Limited time influences individuals to make irrational decisions. When the time needed to make a rational choice (e.g. analyzing pay-offs) exceeds the time available, rationality is bounded (Simon, 1955).

2.3 Privacy loss compensation

The current trend in which large multinationals and even fresh start-ups companies gather more data from their users is not a trend with only negative consequences. According to Jaron Lanier (2013) this development is the fundament of a solution for an even larger problem. Lanier states that the exploitation is not to be stopped; therefore a mutual beneficial relationship should be realized (Leibnizian perspective). The gathering of personal information by firms only formally profits the firms (financially). The users

of the particular information services only profit in informal manners. The firms exploit their personal information by extracting value from the information and subsequently commercializing it. The users have insufficient transparency about these operations and only receive free information services in return. In other words, a formal economy and informal economy are melted in one principal, which only benefits the firms. Clients of the services are persuaded to the services with 'free candy' like a social networks or personalized search results. The information service is not free. The users pay with their personal information (privacy), often without being aware of it. Instead of discussing 'what to share' and 'how to protect' the topic of discussion should be 'how to reward sharing'. This change is fundamental to reaffirm the popular used prophecy 'information is the new oil'.

For several decennia new technological innovations replace the jobs of employees to gain efficiency advantages. The urge to optimize operations effects the financial position of firms and clients (a snake biting its own tail principle). The replacement of jobs by automation or for instance the rise of 3d-printers shrinks the available jobs, while the worldwide population continuous to grow. The working middle class in a country is said to be the largest determinant for a healthy economy. By the urge of efficiency and the growth of automation the financial position of the middle class is endangered, which subsequently endangers the economic stability. A decreasing financial position for the middle class results in shrinkage of markets for businesses. This consequently endangers the businesses and its employees.

To reaffirm the prophecy *'information is the new oil'* the fundamental transactions between firms and its clients needs to become more equal by valuing information for the customer as well. The rewarding of the customers is in need to be formalized. The reconsideration of the value of personal information sharing will in the long term strengthen (or consolidate) the financial positions of the middle class. The financial reward for information can be used to maintain the profitable relationship between the firm and the customer.
A company called Datacoup anticipated this trend by starting an information market-place. The company's business model applies the value-for-information principle. The firm offers a financial reward for connecting services that share your personal information. Social media information, credit card information, health information and location information can be shared with the company.

Connecting more services equals higher rewards. In this way the company benefits by exploiting your personal information and the individual benefits with a financial reward. This redefines the fundaments of business models of information-based services.

This individual case applies the basics of the information-valuing-principle in an actual information market. Application of the principle in other branches like social media-, governmental- or insurance organizations create more opportunities to review the perspective on the information economy. Monetizing personal information creates a sustainable informational economy in the long run.

3 RESEARCH DESIGN

This chapter contains seven sections which clarifies the research design. Firstly the purpose and goals of the research are presented; secondly the research model is presented. Section 3 sums up the research questions accompanied by the hypotheses. Finally the procedure, data collection and data analysis methods are explained.

3.1 Purpose and goals

The purpose of this thesis research is threefold:

1. *Create statistical empirical evidence which illustrates the competence levels that users of Facebook hold regarding their ability to estimate the privacy risks, benefits and trust they are exposed to or uphold and therewith illustrate their ability to accurately perform their privacy calculus.*
2. *Explore and validate the suggested competence model to predict the variables of the privacy calculus.*
3. *Empower the users of SNS with control and awareness in their decision to disclose personal information by creating applicable recommendations for providers, governments and users to improve their behavior regarding performing the privacy calculus.*

3.2 Research model

The research of Krasnova & Veltri (2010) fundaments the conceptual model of this thesis research, however this research is designed with entirely different goals. In this research the **estimation competences**, **skills** and **knowledge** of users are analyzed. The estimation competences are the ability to estimate and recognize certain variables in personal disclosure decisions. The skills are the abilities to sufficiently execute and understand a specific privacy related task on Facebook (Privacy Policy and Privacy Control). The knowledge is the information and awareness that the users have regarding data handling operations of information services. The 'willingness to disclose personal information' is deliberately disregarded in this research since the focus is on competences, skills and knowledge. Image 3.1 illustrates the academic origins from the variables.

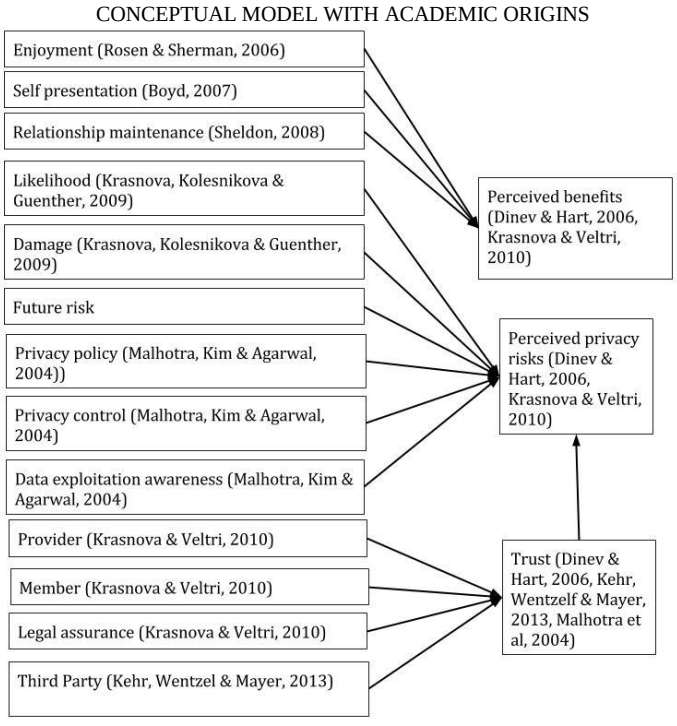

FIGURE 3.1

Perceived privacy risks variables
The estimation competences analyzed with the privacy risks are perceived likelihood (PL), perceived damage (PD) and perceived future risks (FR). The variables *likelihood* and *damage* are more logical than the FR variable. The FR variable is included in the research since the long term consequence of disclosure decisions are of great importance since the sector is subject to rapid changes. New services, new technologies and new legislation continuously arise, the effect of these changes are unclear and hard to estimate. The consequences on the long term are also subject to bounded rationality. The time scope between the disclosure of personal information and the consequences of it is extensive (p. 12). These arguments make the FR variable an interesting topic of research. The PD, PL and FR are variables that are assumed to predict the perceived privacy risks according to Krasnova & Veltri (2010).

In this research PD is the financial, reputational or psychological negative effects inflicted by a privacy risk situation. PL is the probability of actual occurrence of a certain privacy risk in the present. The timeframe of the PL and the FR is the main difference between the two variables. The FR is the probability of the actual occurrence of a certain privacy risk over a longer period of time (> one year) and with greater uncertainties. The estimation of the likelihood is split up in two variables because users are assumed to be able to estimate the likelihood of occurrence in present situations better than in long term situations. The swiftly changing technology sector causes this.

The IUIPC model by Malhotra, Kim and Agarwal (2004) is used to structure two specific skills and the knowledge variable in this research. The original concepts of *collection*, *control* and *awareness* are in this research the Privacy Policy (PP), Privacy Control (PC) and Awareness of data exploitation (ADE). The first two are specific skills and the latter is the knowledge variable in this research.

The PP-skill is defined as the ability to which extent users understand and interpret the data collection statements (privacy policies) of Facebook. The PC-skill is the ability to manage / control the privacy (risks) on the platform. The ADE-knowledge variable is the level of consciousness regarding the use of personal information for commercial purposes.

Perceived benefits variables
In this research three variables to predict the perceived benefits are included, all of them are estimation competences. The Perceived Enjoyment (PE) in this research is the amount of pleasure or joy that users gain by disclosure of personal information. The Perceived Self-presentation (PS) is the amount of favorable representations of the individual that is the result of personal information disclosure. The Perceived Relationship Maintenance (PR) is the amount of upkeep that is executed to maintain or improve relationships with other individuals.

Perceived trust variables
The dimensions of trust are based on actors that are assumed to influence the perceived trust. The Perceived Provider Trust (PPT) in this research is the trust in the service

supplier of the social network platform, in other words Facebook. The Perceived Members Trust (PMT) is the trust in the participants of Facebook. In this case all Facebook account owners. The Perceived Legal Trust (PLT) is the trust in parties or activities of independent parties (like legislation) to protect the users from privacy violations. In this research the legal assurance provider is the Dutch government. The Perceived Third Party Trust in this research is the trust in all partner firms with whom Facebook shares information.
The dimensions PPT and PMT are mentioned as predictors of *Perceived Trust* in the literature research chapter 2. The dimensions PLT and PTPT are added in this research since they are hypothesized to predict the PT as well. All four dimensions are used to measure the estimation competences.

3.3 Research questions & hypotheses

It is unclear if users of SNS possess sufficient competences, skills and knowledge to accurately perform the privacy calculus. The following research question aims to uncover this obscurity:

The main research question:

1. **To which extent are Dutch adult Facebook users able to accurately perform the privacy calculus?**

Sub-questions:

1. To which extent are Dutch adult users of Facebook able to accurately estimate the privacy risks they are exposed to when disclosing personal information?
    a. To which extent do Dutch adult Facebook users possess the skills to decide to disclose personal information?
    b. To which extent do Dutch adult Facebook users have knowledge to decide to disclose personal information?
2. To which extent are Dutch adult Facebook users able to accurately estimate the benefits they experience when disclosing personal information?
3. To which extent are Dutch adult users of Facebook able to accurately estimate the level of trust?

Hypotheses:

1. Most users do not possess a high competence level to accurately estimate the privacy risks they are exposed to when disclosing personal information.
    a) Most users have low skill levels to decide to disclose personal information.
    b) Most users have low knowledge levels to decide to disclose personal information.
2. Most users of Facebook have a high competence level to accurately estimate the perceived benefits of disclosing personal information.
3. Most users of Facebook have low competences in estimating the level of trust in the platform entities.

3.4 Sample

Facebook is the largest online social network currently available. The original website was launched in February 2004 by the current CEO Mark Zuckerberg. The website was intended for Harvard college students, however other universities started using Facebook as well. The platform went public two years after its founding. In May 2012 the company filed for an initial public offering (IPO) on the NASDAQ. The firm was valued at $104 billion, which is considered as the largest valuation for a company that went public for the first time. The IPO raised $ 16 billion, which is the third largest in the history of the United States. The total number of monthly active users in 2013 tops at 1.230 million. The monthly growth among users in the same year decreased to 0.97% worldwide. As of September 2013 71% of the online American adults use Facebook. This number makes Facebook the most broadly adopted social network (Instagram 17%, Pinterest 21% and LinkedIn 22%[3]. The Dutch have 9 million registered accounts, of which 8.1 million are adults (Facebook cijfers juli 2013, 2013).

Due to the large the user base of Facebook is diverse. Facebook is popular across a variety of demographic groups and therefore has no typical user group. Other social networking sites have more unique demographic user profiles. Pinterest attracts for instance mostly female users (female / male ratio: 4/1). LinkedIn is popular among graduate students and internet users from higher income households. The typical users of Twitter and Instagram are younger adults, urban citizens and non-whites (Social Network Fact Sheet, 2013).

A characteristic by Facebook which stands out from other social networks is the high level of user engagement. Of all Facebook users 63% of the users visit the website or application at least once a day. A great deal of users visit the website or application even multiple times a day (40% of all active users). Instagram also has high user engagement numbers, 57% visits once a day and 35% visits multiple times a day. The user engagement of Twitter is remarkably smaller than Facebook's user engagement. Twitter's user engagement reaches only 46% for visiting once a day and 29% multiple times a day (Social Networking sites and our lives, 2011).

The population of this research consists of Dutch Facebook users between the ages of 18 and 65. The minimum age of 18 is chosen since this is the age when individuals become adults in the Netherlands. Rationality and awareness of decision making are important factors in this research and therefore the age of adulthood is chosen. Children and adolescents are interesting groups to do disclosure research on, however these groups are excluded in this research. Conclusions for the sample are not directly useful for adolescents, although they give an indication.

The maximum age of 65 is chosen since statistical evidence about the age of the users of Facebook indicates that <40% of the group has a Facebook account[2]. The threshold to partake in this research is assumed to be too high. This is not a representable number of the group anymore. The research prefers a variety of educational backgrounds, because

[3] http://www.socialbakers.com

the level of education seems to impact decision making. A balance between male and female is wanted, Facebook has a nearly equal balance between male and female users (male 49%, female 51%). A requirement is that the participants use Facebook as an online social network, the chosen methodology incorporates features from Facebook that require a minimum entry-level.
Dutch users of Facebook are selected for this research. The Dutch have several interesting characteristics about their online disclosure behavior and general internet usage. The percentage of people that are online tops at 93% (5th in Europe). More than half of the Dutch internet users have a Facebook account (53.5%); this is medium / high for European standards. In the United Stated this percentage is 43%. Facebook is the most used social media in The Netherlands, next are Twitter (12%), LinkedIn (9%) and Google+ (8%). Of all individuals that are online 78% uses social media.

The age group of the Dutch Facebook users slightly differs from the worldwide age groups. The group of 25-34 is on average larger in the world than in The Netherlands. While the group of users with ages smaller than 18 is smaller in the world (Table 3.4).

AGE DISTRIBUTION

Age group	Percentage of Facebook users
13-15	4.6%
16-17	4.8%
18-24	18.5%
25-34	21.6%
35-44	17.9%
45-54	16.1%
55-64	10.0%
65-100	6.5%[4]

TABLE 3.4

Non-users of Facebook are not incorporated in the scope of this research. Recent studies by Govani & Pashley (2007) and Acquisti & Gross (2006) do not agree on the fact if non-users of Facebook do not participate in the network because of their privacy concerns. Govani and Pashley state that the decisions to participate in the network are due to openness of the provider and publicly posting personal information. However, Acquisti & Gross (2006) concluded that the concerns for privacy are not the reason not to join because many users who did join had similar beliefs. Facebook is subject to continuous changes in policy and features and users who opted out on the service are therefore not aligned with the target group of the research. There is no need for prior knowledge about privacy risks and information disclosure.

3.5 Procedure / methodology

This research consists of five parts. First a literature study was executed to explore the concepts and get familiar with the research topics. This phase originated the research design and the conceptual model. Next on was the development of the survey. This was executed with previously collected concepts from the literature. The survey was distributed among the research sample to collect data. Parallel the data gathering of the expert panel ran. After closing the survey the data analysis phase started. Linear regressions analyses are executed to corroborate the conceptual model. This resulted in a new model. In this phase the data gathered from experts is analyzed, which resulted in a golden standard. Prediction models and a prototype of a tool (Appendix K) are

[4] http://www.socialbakers.com/facebook-statistics/netherlands

constructed with the output of the regression analysis. Finally the users are categorized in scales to illustrate competences, skills and knowledge.

Literature research
The main aim of this part of the research was to structure the fundaments of this research. The research questions, hypotheses, concepts and theoretical model are results of this phase. Two methods were used to systematically research literature: snowballing and the forward method. The snowballing method is performed to result in old and new research on the selected topics. The main focus of this method was to explore the theoretical framework and get familiar with the body of knowledge regarding the research topics. The forward method is conducted to explore the latest and groundbreaking research on the topic. This guaranteed the originality of this research.

Instrumentation
The main research instrument used for data collection is an online survey. The survey consists of 10 parts with a total of 48 questions. The first part of the survey collects demographics, while the other 9 parts gather data about the estimation competences, skills and knowledge. A distinction is made between questions that ask to score the perceived value of elements (LIKERT scale) and skills questions (Boolean statements). Competences of estimating and recognition are easier measured via LIKERT scales, while specific knowledge and skills are measured with dichotomous questions. Multiple variables are combined in one set of questions to limited the duration of the survey.
Several questions per part aim to explain a construct. A scoring on just one item would not represent the total construct. A summation of the item scores would not capture the essence of the total construct. Therefore mean scores are composed with the individual item scores.

The demographical questions in the survey gather information about whether the respondents match the research population. Respondents who do not match the research population are excluded from the research.
The survey contains ten text-written hypothetical situations. The hypothetical situations are constructed in a realistic manner. In appendix D a table is presented which clarifies methods per variable. The hypothetical situations are derived from stories that actually occurred. The frequency of occurrence of the situations is unclear. The situations describe general situations; specifics are left out to make sure the respondent is able to identify himself with it. The questions contain words like ‘you’ and ‘yours’ to increases the ability to imagine to the situations better.

Part 3 and part 5 contain true or false statements to measure the skills of the respondents regarding the research topics. The option ‘I do not know’ is added to prevent individuals from guessing. In the analysis the ‘I do not know’ option is interpreted as if the question was wrongly answered. The statements selected for part 3 of the survey have the characteristics that they are most important, basic and widely discussed. The questions and the statements are structured in a similar way by which they are presented in the privacy policy of Facebook. Part 5 contains questions regarding the topic of privacy control on Facebook. The selected situations describe five (possible) functionalities which improve the control over privacy for the users. The users are asked whether the

functionalities exist or not. The knowledge of existence of the functionalities indicates whether the users have control over their personal data.

The variable in part 4 measures the knowledge which Dutch Facebook users have regarding the awareness of the data exploitation. Manipulated situations which occur on a daily basis for the respondents are used to measure this knowledge. The selected situations are structured in a recognizable manner for the respondents.

The trust statements from part 7 are derived (and adjusted) from the research of Krasnova & Veltri (2010). Part 8 and 9 again show images of typical Facebook posts. The participants are asked to estimate the particular variables.

3.6 Data collection

The data collection in this research consists of collection from experts and Facebook users. All the participants in this research are approached in person and originate from the personal network of the researcher. The personal approach is a deliberate choice in this research since the duration to fill the survey is exceeding 20 minutes. This method is used to reduce the risk of having too little participants. The personal network of the researcher is addressed to increase the chances of sufficient participants even more. These choices resulted in a low drop-out ratio (8.7%) and sufficient number of respondents (N=123). The research is anonymously conducted. The users are informed of the anonymity beforehand.

The data gathered from users of Facebook is compared with data gathered from experts. Experts are asked to fill exactly the same survey. The results of several experts are combined to form a panel score. If the panel score meets certain requirements the score is labeled as golden standard. Since the golden standards are structured with the objective interpretations of the opinions of experts several criteria are applied to improve the usability (Appendix E). When the score of the expert does not match with the criteria the score is not used to structure a golden standard. The experts are influenced by bounded rationality as well. Lack of information or decisions on an automatic pilot influence the users as well as the experts. For this reason scores of multiple experts are used to construct the golden standard. In this research the results of the competence analysis are presented with and without the excluded experts to prevent expert selection bias. The weight of the experts is assumed to be equal since their knowledge is not quantifiable.

Experts that have demonstrable theoretical and applicable knowledge about a certain topic or variable are approached.
The expertise of the expert is checked in a meeting. The experts are desired to have a Facebook account; this is not an obligation to the research. Age, gender, origin and language are neither requirements. The experts are not mentioned with their names in this report to secure the anonymity commitment.
In the meeting with the expert he/ she is asked to answer questions that have the goal to express the sentiment of the expert. A critical set of skills and knowledge about the variable are expected from the expert. The age, origin and for instance the political preference are possible variables that could influence the expert's opinion.

The set of criteria to decide whether to use the panel score in the competence analysis resulted in a major effect to the research. For the variable and factors of perceived trust two experts were inputs for the panel score. The deviation between the two experts exceeded the maximum. This resulted in the exclusion on the competence analysis on the variable perceived trust. The criteria for the perceived benefits and perceived privacy risks are obtained. These variables are included in the competence analysis. Analysis shows that Expert 2 possess an extremely deviating (and critical) opinion on several variables (Appendix J).
The exclusion of the panel score on the variable perceived trust decreases the validity of the golden standard. Although perceived trust is said to be the mediating factor between perceived risks and benefits, the competence analysis of estimating the perceived trust is not used in the conclusion of the research. The scores with Expert 2 included have insufficient validity but are included for comparison purposes.
The experts were not asked to answer the dichotomous questions since these questions do not require a golden standard. These questions are either correct or incorrect.

3.7 Data analysis

The survey consists of four main parts: demographics, perceived privacy risks, perceived trust and perceived benefits. Each of the parts contains variables and items that are analyzed for inter-correlations. The parts of the survey and the associated questions can be found in Appendix D.

The complete data set is screened and trimmed before the statistical analysis start. Screening the data set improves the quality of the sample. To prevent biased case deletion the selection criteria and process of data screening is elaborated in Appendix B. In total n=7 is deleted from the research sample. In total 8 scores of 7 different questions are excluded for statistical analysis.

The analysis of the demographic items results in how the sample is distributed among age, education, gender and the importance of online privacy.

Combining scorings from the questions result in comparing group means. Some groups are clear cut by definition, like gender (males and females). Other groups are created by cut-off points in the distributions. The means of the groups are compared with the Independent T Test and the Mann-Whitley U test. Clustering the sample in for example an ‘old age group’ and a ‘young age group’ opens up the possibility to analyze difference between the two groups.

Model validation

Analysis of the variables consist of four parts. The first part of the analysis is executed to check the consistency of the items (Cron bach’s Alpha), next on the descriptive and frequency analysis are performed to roughly indicate the results with face-validity as a result.
The second part consists of Factor analysis (Principal Component Analysis). The third part of the analysis of variables contains analysis for inter-correlations between the factors. The inter-correlations are analyzed with Linear Regression Analysis (in some

cases Kendall Tau C or Spearman Rho). The aim of this step in the procedure is to corroborate the model. This is executed by testing correlations between the predictor variables and the main variables.

Factor analysis

The research started with thirteen different factors that are assumed to determine (or influence) the three overlaying variables. The goal of factor analysis at this stage in the research is to confirm the independence of the chosen factors. The focus of the PCA is based on possible factor reduction to move to a non-redundant and efficient validated model. Inter-correlations and overlapping factors are reduced from a large set of factors to a smaller set of factors.

The factors are addressed as separate concepts from the start of the PCA. The procedure was executed as following (full procedure is included in Appendix C):

All the factors that are suggested to be predictors of the overlaying variables are input for the PCA, the structure of the PCA is illustrated in table 3.7. Direct oblimin rotation is the chosen rotational method to increase interpretability of the components. This method beholds the assumption that factors are allowed to be correlated internally.

The factor analysis was executed on the following factors:

FACTOR ANALYSIS STRUCTURE

Overlaying variable	Input for PCA:
Factor analysis Perceived Benefits	Perceived enjoyment, Perceived Self-presentation and Perceived Relationship Maintenance
Factor analysis Perceived Privacy Risks	Perceived damage, Perceived likelihood, Perceived Future Risk, Privacy Policy, Exploitation awareness and Privacy control.
Factor analysis Perceived Trust	Perceived Provider trust, Perceived users trust, Perceived legal assurance trust and Perceived Third party trust

TABLE 3.7

The interpretation to the PCA is key to determine the number of derived components. The interpretation of just one statistic is not sufficient to determine the number of components. The interpretation of the Kaiser-Meyer-Olking Measure of Sampling Adequacy (KMO), Percentage of Eigenvalue explained, Scree plotting and analysis of Components and Patterns is used to decide on the number of overlaying components.
To check whether the results of the PCA are useful regression analysis with and without the new components are run. When the correlations improve (and the components still make sense) the component replaces the original variables. Otherwise the original model is preserved.

Research by Krasnova & Veltri (2010) states correlations between the variables perceived trust, perceived benefits and perceived privacy risks. Inter-correlations between the variables are measured as well. The statistical analysis used to perform this measured is Linear Regression Analysis.

Expert score analysis
Users are categorized in competence levels to illustrate their abilities. To categorize the competence levels of the users their scores are compared with the golden standard. The deviation between the users and the experts indicate the competence per measured factor / variable. The number of experts and the mean scores of the experts are included in Appendix E. The scores of the experts per factor and per variable are combined in a mean score. The minimum, maximum, mean and standard deviation are analyzed to determine if the mean score can be used in the competence analysis with the users.

Expert vs. user analysis
The competence analysis results in deviation scores between the golden standard and the users. The competence analysis requires two elements before the analysis can run. The first element is the golden standard per factor / variable (μe). The second element is the mean score of the users per factor / variable(μu). Δ (i) is the difference per item. The difference between the mean score of the experts and the users is calculated for all items.

$$\Delta(\boldsymbol{i}) = \boldsymbol{\mu e} - \boldsymbol{\mu u}$$

Scale determinations
Scales to categorize the users in competence level are determined for Likert scale items and dichotomous questions. The competence scale for Likert scale items is based on the earlier mentioned Δ (i) and the competence scale for dichotomous questions is based on the number of correct answers. Method one uses the deviations between the experts and the users, while method two does the same but corrects for the prediction values in the model.

Each scale is developed with five competence levels (Competence level 1 till Competence level 5). Competence level 1 represents low competence and competence level 5 represents high competence. The size of the deviations determines the competence level. The different methods and different measurements require specific criteria. The calculations of the competence scales are included in Appendix H. The three main methods that are applied to construct the scales are elaborated there.
The next step in the analysis is the categorization of the users in the competence levels.

Hypotheses testing
In the last part of the statistical analysis the hypotheses are tested. To test the hypotheses the estimation competence levels, the skill levels and the level of knowledge are required. The categorization in estimation competence levels is run in three different ways. The discrepancy between the methods is based on whether the method uses corrected expert scores and / or corrects for trust.

1. The categorization is executed without expert 2, and therefore the inter-correlation with *Perceived Trust* is disregarded
2. The categorization is executed with expert 2, however the inter-correlation with Perceived Trust is disregarded.

3. The categorization is executed with expert 2 in the golden standard. The Perceived Trust inter-correlation with Perceived Privacy Risks is included in this analysis.

4 RESEARCH RESULTS

In this chapter three different research results are presented. First the **demographic characteristics of the research sample** are presented. The next section is used to present results from the **model construction analysis**. In this part a new model is presented which illustrates the correlations. Prediction functions are presented which predict the value of the main variables. The last section of this chapter presents the **competence, skill and knowledge levels** of the users of Facebook. In this section the results which are used for hypotheses testing are enumerated.

4.1 Demographic results

In this part of the research results the findings of the demographic questions are presented.

Age distribution

The Facebook users in the research sample are distributed between the range of 18 and 35 years old. The research sample contains little users above the age of 35 as table 4.1 illustrates. Users below the age of 18 are deleted from the study (n=5). No users above the age of 65 are present in the sample.

Age distribution

	Frequency	Percent
18 - 25	62	56.4
26 - 35	33	30.0
36 - 45	3	2.7
46 - 55	7	6.4
56 - 65	5	4.5
Total	110	100.0

TABLE 4.1

Gender distribution

The research sample contains 72 males and 38 females. The gender distribution of the sample is unbalanced. The researcher unintentionally approached more males than females to be potential respondents for the survey. The ratio males / females does not match the ratio of the entire population. No evidence was found that males are more eager to participate in this research.

Level of education

The research sample contains more high educated Facebook users than low educated Facebook users (illustrated in table 4.2). The researcher unintentionally approached more high educated users than low educated users.

Facebook users with a HBO or WO education are labeled as high educated. The sample does not contain users who did not have any education. The level of education of the sample deviates from the population.

Distribution of education

	Frequency	Percent
Lager beroepsonderwijs of voorbereidend middelbaar beroepsonderwijs (LBO of VMBO)	6	5.5
Middelbaar beroepsonderwijs (MBO)	23	20.9
Hoger algemeen voortgezet onderwijs of voorbereidend wetenschappelijk onderwijs (HAVO of VWO)	9	8.2
Hoger beroepsonderwijs (HBO)	46	41.8
Wetenschappelijk onderwijs (WO)	26	23.6
Total	110	100.0

TABLE 4.2

The population consists of more spread education levels.

Visiting frequency

The frequency of visiting the website of Facebook or Facebook apps indicate the integration of Facebook in daily life. Table 4.3 illustrates that the greater part of users visits Facebook several times a day. This behavior aligns to the population behavior.

Distribution of visiting behavior

		Frequency	Percent
Valid	Nooit	2	1.8
	1x per maand	3	2.7
	1x per week	3	2.7
	Meerdere keren per week	10	9.1
	1x per dag	11	10.0
	Meerdere keren per dag	81	73.6
	Total	110	100.0

TABLE 4.3

Frequency of sharing

Most users share little on Facebook (Table 4.4). Combining this statistic with the frequency of visiting Facebook indicates that many users mostly use Facebook to lurk.

Distribution of sharing behavior

		Frequency	Percent
Valid	Nooit	11	10.0
	1x per maand	38	34.5
	1x per week	25	22.7
	Meerdere keren per week	22	20.0
	1x per dag	4	3.6
	Meerdere keren per dag	10	9.1
	Total	110	100.0

TABLE 4.4

Main reason to use Facebook

The main reason to use Facebook is to stay informed (53.6%), which aligns with the frequency of sharing. A little more than a quarter of the users indicated that their main usage of Facebook is to keep in touch with friends and family (28.2%).

Four users choose the option 'Different' and wrote answers that better sooth their opinions. Three of those users wrote that they wanted to select two options because they are equal reasons to use Facebook. One user wrote 'to gather information for my hobby.'

Estimations of knowledge and capabilities

Users were asked to which extent they think that they have knowledge and are capable to the subject of use and misuse of their personal data. Table 4.5 indicates that most users estimate their knowledge and capabilities regarding the topic >5. Most users think that they have above average knowledge and capabilities which can have two explanations. One, users are actually skilled or two users overestimate their knowledge and capabilities. Both results lead to interesting conclusions about the sample.

Distribution of knowledge estimation

		Frequency	Percent
Valid	1	2	1.8
	2	4	3.6
	3	11	10.0
	4	21	19.1
	5	39	35.5
	6	29	26.4
	7	4	3.6
	Total	110	100.0

TABLE 4.5

Importance of online privacy
Users were asked to which extent they think online privacy is important. None of the users indicated that they did not think online privacy was important (score 1 or 2). Most users (82.7%) think online privacy is important (5,6,7).

Privacy policy readings
'Did you read the privacy policy?' Nine of the 110 users entirely read the privacy policy. Most did not read the privacy policy, while 39.1% of the users partly read the privacy policy. This statistic does not illustrate why users only read the privacy policy partly. It is possible that the users who partly read the privacy policy only filtered out the important (relevant) parts or that they stopped reading because of the amount.

4.2 Combined demographic results

The results of the individual demographic questions are analyzed and combined with other questions. Combination these results lead to interesting conclusions.

Older users of Facebook do think that online privacy is more important than younger users do (sig 0.0028). High educated individuals think that online privacy is more important than low educated users do (sig 0.018). This result does not imply that users who think online privacy is important get higher educations.

High educated users do not think that they have more knowledge and capabilities than low educated users. It is possible that users who are highly educated are more aware of the many possibilities of use and misuse of personal data and therefore do not estimate their knowledge and capabilities that high.

A discrepancy exists between the frequency of visiting Facebook and the frequency of actually sharing personal information. Visiting Facebook is almost fully integrated in daily life while sharing personal information stays behind. There is a positive correlation between the frequency of visit and the frequency of sharing.

Statistical conclusions about users who read and users who did not read the privacy policy also give interesting insights. The group of users who read the privacy policy estimate their knowledge and capabilities regarding the use and misuse of personal information higher than users who did not read the privacy policy. The data does not show if the users who read the policy started feeling more knowledgeable or that knowledgeable and capable feeling users started reading the policy.
A similar significance is measured regarding the importance of online privacy and reading the privacy policy. It is not clear if users who read the policy started estimating the importance of privacy higher or that users who estimate online privacy as important started reading the policy.

Users were asked specific questions about the privacy policy of Facebook. Remarkably the users who read the privacy policy do not significantly answer these questions better. It is possible users forgot the content of the policy or are unable to apply it because of the complexity.

Users with high educational backgrounds estimate the likelihood of occurrence of a privacy risk higher than users with low educational backgrounds. Assumed is that these users are more informed about the possible violations.

Results show that the least trusted entity is the third parties. This is a logical result, since these parties do not offer any transparency or control features for the users. And yet are able to access your personal information. The trust in the legal assurance is the largest. This is explained by the fact that the legal assurance has only intentions of protecting users and not misusing their data.

A situation in which an individual shares his phone number on Facebook is perceived as the most risky. Sharing an image of an individual who drinks alcohol in a party outfit is also perceived as a substantial privacy risk.
The most enjoyment is perceived in the situation in which an individual shares she graduated her education. Users that share more personal information perceive more enjoyment in the sharing of others.

4.3 Model construction

The PCA and regression analysis discovered the best fitting model to perform the privacy calculus. The new model is presented below and the adjustments are discussed.
The factor analysis on the factors of the three main variables did not resulted in changes in the model. The original model, consisting of thirteen factors that influence the three main variables is replaced with a more efficient and reduced model. The image (Figure 4.3) below illustrates the new model.

In the model (Figure 4.3) below the correlation between one individual predictor and the overlaying main variable are illustrated. The correlations in the model do not take the inter-correlation and multicollinearity between the predictor variables in account. The inter-correlation and multicollinearity are of little influence on the model.

As a result of analyzing the correlation between the predictor variables and the main variables one variable is deleted from the model. *The perceived provider trust (PPT)* is deleted from the model, since no significant correlation was found. The variables PE, PS and PR remain.
The factors which influence the variable perceived privacy risk (PL, PD and FR) also remain in the model. No new components were structured with the PCA.
The skill variables PP and PC and knowledge variable ADE are not included in the model, since these are not predictors.

INDIVIDUAL CORRELATION MODEL

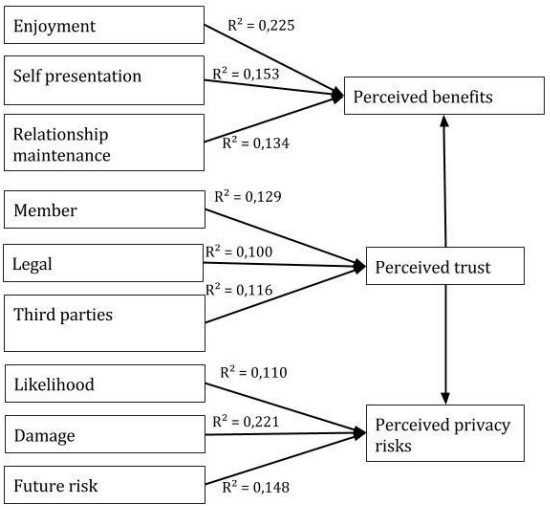

FIGURE 4.3

Interpretation of R^2

Most of the explained variances between the independent variables (IV) and the dependent variable (DV) are interpreted as weak correlations. Table 4.3 illustrates the ranges that determine the interpretations of R^2.

CORRELATION INTERPREATION TABLE

R^2 strength	Interpretation
0 – 0.2	Weak / Slight
0.2 – 0.4	Mild / Modest
0.4 – 0.6	Moderate
0.6 – 0.8	Moderately strong
0.8 – 1.0	Strong

TABLE 4.3

The correlations of *PE* and *PD* are therefore interpreted as mild / modest. The other correlations are present, however interpreted as weak / slight.

For the *PE (R^2 0,225)* the explained variance score is relatively high compared to the other two predictors. It is expected that *PE* scores higher than *PS* and *PR* since the concept of enjoyment and benefits are fundamentally more similar.

The expected results for the predictor variables of Privacy Risks slightly deviate from the actual results. Correlations between *PD*, *PL* and *FR* are present, but were assumed to have higher coefficients of determination (R^2).

The results of the analysis verify for three of the four predictor variables for trust significant correlations. The *PPT* is excluded from the model (Figure 4.3). This result does not align with the expectations. Research by Krasnova & Veltri indicate *PPT* and *PMT* to be predictors. An explanation is given on page 34. Positive correlations between the DV and *Member Trust, Legal Trust and Third Party Trust* are discovered, although the correlations are interpreted as weak.

Model selection

In figures 4.4 and 4.5 the correlations between individual IV's and the DV are illustrated. The correlations coefficients start to deviate when the multiple IV's are the input for the model. Negligible deviations between the correlations occur since multicollinairty, inter-correlations and an intercept are included.

COMBINED CORRELATION MODEL

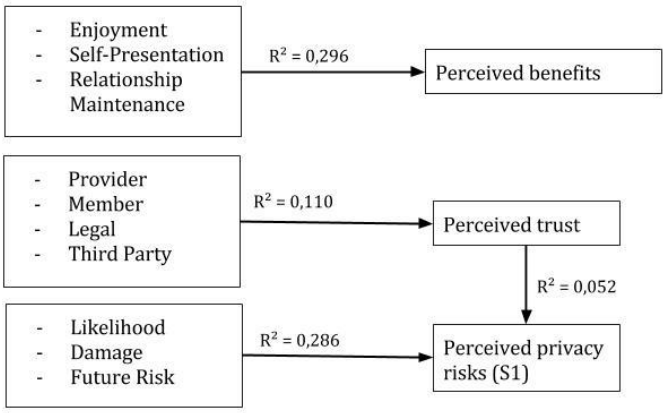

FIGURE 4.4

The goal is the construct prediction models. A prediction model can only contain significant (Beta's) variables. The validity and usability of the prediction model would otherwise be low. On the other hand the chosen model should be as complete as possible, therefore two models are compared.

Two different models are presented to demonstrate the best fitting model.

- Model 4.4 is constructed with all predictor variables (only significant correlations).
- Model 4.5 is constructed with significant Beta coefficients taking in account (and significant correlations)

A correlation between Perceived Trust and Perceived Benefits is not discovered. The discovered correlation between PT and PPR is weak. This result aligns with the expectations of Dinev & Hart (2005).

Comparisons between model 4.4 and 4.5 illustrate a slight improvement. The R^2 values for the prediction of the Perceived Trust and Perceived Privacy Risks improve, while the R^2 for the Perceived Benefits slightly decreases.

CORRELATION MODEL WITH SIGNIFICANT BETA'S

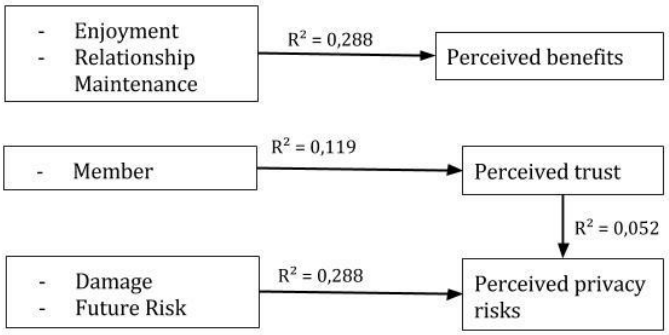

FIGURE 4.5

The excluded variables in model 4.5 (compared with model 4.4) are the PS, PPT, PLT, PTT and PL.

The best fitting model is model 4.5.
In the next section the prediction functions according to model 4.5 are presented. The construction of the prediction functions is ordered according to the main variables. The tested models and causes of variable exclusions are presented there.

Perceived benefits model

Model 4.5 (page 28) demonstrates that the variables *PE* and *PS* have significant correlations and significant Beta coefficients to be used in a prediction model for the *Perceived Benefits.* In this section the functions to predict the main variable are constructed and the exclusion of the *PS* variable is elaborated. The function is used to predict the estimation competences of the users and fundaments the prototype of a tool.

Benefit Model Summary

Model	R	R²	Adj R²	Std. Error of the Estimate
1	.505[a]	.255	.248	.92712
2	.537[b]	.288	.275	.91047

TABLE 4.4

Benefit Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.553	.574		.963	.338
	PE	.654	.107	.505	6.083	.000
2	(Constant)	.019	.612		.031	.975
	PE	.553	.115	.427	4.820	.000
	PR	.235	.105	.198	2.233	.028

Table 4.5

A Stepwise Linear Regression Analysis is used to determine the best fitting model. The analysis constructed two models with different predictor variables. In none of the models the PS variable is included because the Beta coefficient of this variable is not significant. The included variables are presented in Table 4.5

The R^2 and R^2_{adj} of model 2 surpass model 1 (See table 4.4). R is stronger and the Std. Error of the Estimate is lower. Model 2 is therefore used to construct the prediction function.

The ANOVA-table which corroborates this statement is included in the Appendix G.

The exclusion of *SP* negatively influences the completeness of the model. The residual predictors explain 28.8% of the variance in the dependent variable. Both coefficients are suitable to act as predictors.

The Scatterplot (Figure 4.6) clearly illustrates the correlation. The graph also shows a few outliers.

SCATTERPLOT BENEFITS

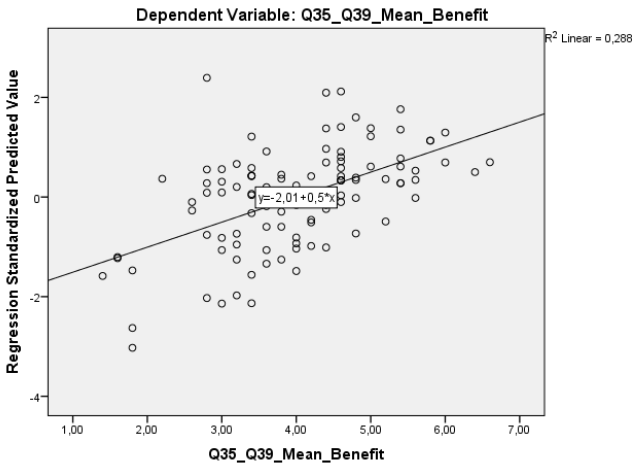

FIGURE 4.6

The basic function to predict Y *is:*
$\hat{y} = \alpha + \beta 1\chi 1 + \beta 2\chi 2$
Constructing the model function to predict the *Perceived Benefits* results in: $\gamma = 0{,}019 + 0{,}553\chi 1 + 0{,}235\chi 2$
Where X1 = User mean score for PE, and X2 = User mean score for PR.
The prediction function is interpreted as usable for the prototype.

Perceived privacy risk prediction model

Model 4.5 demonstrates that the variables *PD* and *FR* have significant correlations and Beta coefficients, these can be used in a prediction model for the Perceived Privacy Risks. In this section the functions to predict the main variable are constructed and the exclusion of the *PL* variable is elaborated. The function is used to predict the estimation competences of the users and fundaments the prototype of a tool.

Privacy Risk Model Summary

Model	R	R^2	Adj R^2	Std. Error of the Estimate
1	.470[a]	.221	.214	.84566
2	.536[b]	.288	.275	.81224

TABLE 4.6

A Stepwise Linear Regression Analysis is used to determine the best fitting model. The analysis constructed two models with different predictor variables. In none of the models the *PL* variable is included because the Beta coefficient of this variable is not significant. This is an unexpected result since the *likelihood* of a privacy risk is assumed to be one of the two main predictors of risks in general, the other one being damage. The included variables are presented in Table 4.7.

Privacy Risk Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.319	.329		10.077	.000
	PD	.383	.069	.470	5.532	.000
2	(Constant)	2.525	.403		6.260	.000
	PD	.319	.070	.391	4.583	.000
	FR	.242	.076	.271	3.173	.002

TABLE 4.7

R^2 and R^2_{adj} for model 2 are stronger (Table 4.6). The Power of R is stronger and the Std. Error of the Estimate is lower for Model 2. Model 2 is therefore used for the prediction model.

The model that is constructed with the calculated coefficients and constant results in:

$$\hat{y} = 2{,}525 + 0{,}319\chi 1 + 0{,}242\chi 2$$

Where X1 = Perceived Damage and X2 = Perceived Future Risk.

Scatterplot (Figure 4.7) illustrates a graph of the function. This graph demonstrates the positive modest correlation. Some outliers from the regression line are present. The model is interpreted as usable for the prototype and the analysis of the estimation competences.

SCATTERPLOT PRIVACY RISKS

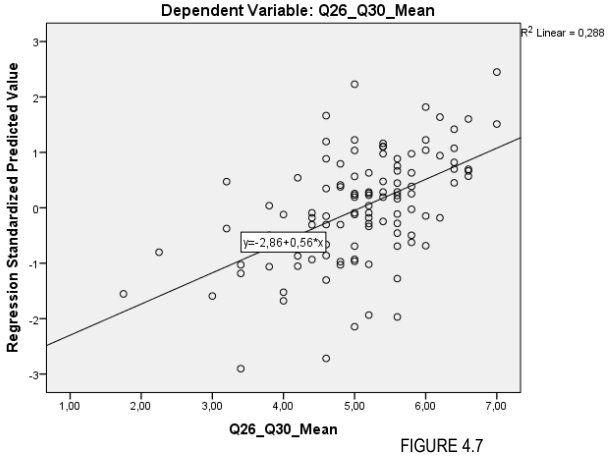

FIGURE 4.7

Perceived trust prediction model

Model 4.5 demonstrates that the variable *PMT is* significantly correlated and has significant Beta coefficients; this variable can be used in the prediction model for the *Perceived Trust*. In this section the function to predict the main variable is constructed and the exclusion of the *PPT, PLT and PTT* variables are elaborated. The function is used to predict the estimation competences of the users and fundaments the prototype of a tool.

TABLE 4.8

Trust Model Summary

Model	R	R ²	Adj R ²	Std. Error of the Estimate
1	.345	.119	.111	1.05669

Trust Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	2.768	.242		11.424	.000
	PMT	.258	.068	.345	3.817	.000

TABLE 4.9

TABLE 4.10

Excluded Trust variables

Model		Beta In	t	Sig.	Partial Correlation
1	PPT	-.033[b]	-.327	.745	-.032
	PLT	.168[b]	1.551	.124	.148
	PTPT	.197[b]	1.778	.078	.169

A Stepwise Linear Regression Analysis is used to determine the best fitting model. The analysis constructed one model (Table 4.8). The model excludes the PPT because of insignificant correlations. The PLT and PTPT variables are excluded since the Beta coefficients of these variables are not significant. This is an unexpected result because the *likelihood* of a privacy risk is assumed to be one of the two main predictors of risks in general, the other one being damage. The included variables are presented in Table 4.9.

The excluded variables (and the insignificant Beta coefficients are illustrated in table 4.10. The model function that is constructed to predict the *Perceived* Trust results in $\hat{y} = 2{,}768 + 0{,}258\chi 1$

Where X1 = PMT mean

Scatterplot 4.8 illustrates the measured correlations and coefficient. The spread variance indicates the low R^2 value. The applicability of the Trust function is therefore interpreted as weak and will not be use for the tool.

SCATTERPLOT TRUST

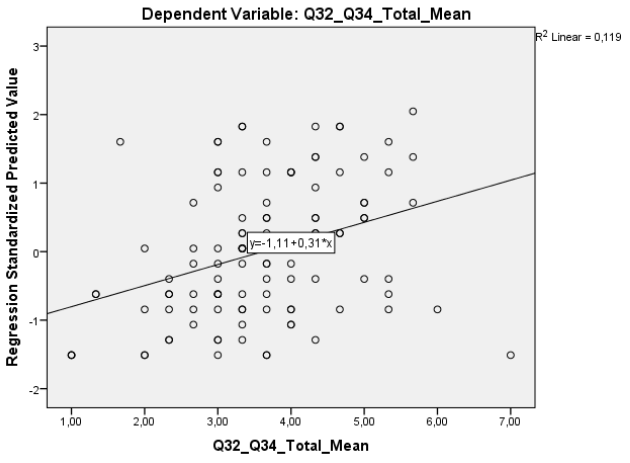

FIGURE 4.8

4.4 Main variable inter-correlations

Inter-correlations between the main variables are also analyzed. The results of this analysis deviate from the expected results. The correlations that are analyzed in this part were *Perceived Trust – Perceived Benefits* and *Perceived Trust – Perceived Privacy Risks.* A positive correlation was expected in the first measurement, while the latter measurement was expected to have a negative correlation.

The analysis discovered only the negative correlation between *Perceived Trust* and *Perceived Privacy Risks*. The power of the correlation and the R^2 are both smaller than expected, therefore the influence of *PT* on *PPR* is interpreted as weak. Dinev & Hart (2005) predicted the correlation, however did not expect it to be that small.

Possible causes for the discrepancy of the correlation of *PT* and *PB* are found in the domain of definitions. Dinev & Hart researched *Internet Trust* while this research analyzed *Perceived Trust in entities*. Another difference is that they found Provider Trust to be strongly correlated, while this research does not.

4.5 Ability analysis

In the expert analysis the results indicated that Expert 2 valued extreme deviating scores compared with the rest of the participants of the expert panel. Two different results are presented in this part of the chapter, results with Expert 2 included and excluded.

Excluding expert 2 from the research results in violating the criteria that the number of experts per variables consists of minimum of 2. Statistical analysis for the trust variable without Expert 2 is therefore not usable. At first all the results of competence levels are presented (except for the trust variables). The large deviation scores on trust by the experts indicate the complexity of estimating the value.

The predicted competences for the *Perceived Benefits* and *Perceived Privacy Risk* are presented at the closing section of this chapter.

Competence scores: Privacy Risk estimation (excluded expert)

Table 4.11 illustrates the competence of estimation of the privacy risk variables. This table illustrates the competence levels of the users when Expert 2 is excluded from the golden standard.

Notably, most users score competence levels ≥ 3 on *PD, PL* and *FR.*

These results indicate that most users have sufficient or high competences in estimating the privacy risks.

Users have high abilities to estimate the *PL* and the *FR*. The estimation of *PD* is sufficient / high. The urgency to create recommendations to increase these competences is therefore absent.

PRIVACY RISK COMPETENCES

	PD	PL	FR
Competence level	Frequency	Frequency	Frequency
1	6 (5.5%)	7 (6.4%)	6 (5.5%)
2	9 (8.2%)	10 (9.1%)	6 (5.5%)
3	30 (27.3%)	10 (9.1%)	15 (13.6%)
4	34 (30.9%)	40 (36.4%)	34 (30.9%)
5	31 (28.2%)	43 (39.1%)	49 (44.5%)
Total	110 (100%)	110 (100%)	110 (100%)

TABLE 4.11

Competence scores: Privacy Risk variables (included expert)

Competence results for the factors of the *Perceived Privacy Risks* when Expert 2 is included differ from the results when Expert 2 is excluded.

Table 4.12 illustrates that the competences in *PD* increase mainly for the highest competence level. The results for the *PL* move towards an average competence level (Level 3). The *FR* competences decrease.

Competences score of Privacy Risk variable				
		PD	PL	FR
Competence level		Frequency	Frequency	Frequency
Valid	1	5 (4.5%)	8 (7.3%)	10 (9.1%)
	2	10 (9.1%)	11 (10.0%)	7 (6.4%)
	3	14 (12.7%)	17 (15.5%)	20 (18.2%)
	4	35 (31.8%)	38 (34.5%)	34 (30.9%)
	5	46 (41.8%)	36 (32.7%)	39 (35.5%)
	Total	110 (100%)	110 (100%)	110 (100%)

TABLE 4.12

Most of the users still score sufficient or high competence levels regarding the estimation of privacy risks. Users have high abilities to estimate the *PL* and the *FR*. The estimation of *PD* is sufficient / high. The urgency to create recommendations to increase these competences is therefore absent.

Hypothesis H1 is therefore rejected. Users do possess a high competence level to accurately estimate the privacy risks they are expose to when disclosing personal information.

PRIVACY POLICY SKILL

Skill results: Privacy Policy

The results of the skill questions about the privacy policy of Facebook contain two remarkably results. Question 11 asked the users if Facebook is the owner of your Intellectual Property (IP) that you put on Facebook. Less than 20% of the users correctly answered this question and know that Facebook does not own your IP. The users continue to use Facebook while presuming this fact. This indicates a certain careless attitude towards the data collection by Facebook.

Question 14 asked the users if they think that Facebook deletes personal identifying items from your data before sharing it with third parties. The result of this question verifies that users do not know the answer. In total 67% answered the questions with 'I do not know', while only 10.9% correctly answered the questions. The users continue to use the service without being certain if Facebook possibly violates the privacy of the individual. These two basic statements supply evidence to think that users have insufficient skills in the privacy policy of Facebook. The attitude to improve this seems to lack.

	PP
Competence level	Frequency
1	67 (60.9%)
2	25 (22.7%)
3	13 (11.8%)
4	5 (4.5%)
5	0 (0%)
Total	110 (100%)

TABLE 4.13

Table 4.13 confirms this evidence. A total of 83.7% of the users have low skill levels regarding their knowledge and interpretation of the privacy policy. These results emphasize the urge to create recommendations to improve this specific skill level.

Skill results: Privacy Control

The skill levels regarding the topic of Privacy Control on Facebook also result in most (65.4%) users that possess insufficient skills. The typical result of question Q21 gives away an indication of the total skill scores of the users. Question 21 asked '*Is it possible to choose the specific recipients of a message when you share via your timeline?'*. It is in fact possible to do this, although the answers to this question indicate that not all users are aware of this (65.6%). Facebook implemented extra options to enhance the (feeling of) control of privacy by developing transparency features. The knowledge and awareness of these concepts are tested in the survey. Question 25 asks *'Does Facebook offer a log which sums up all your Facebook activities?*' 55% of the users wrongly answer this question. In other words, Facebook offers these features to increase the privacy control and trust of the users but the most users are not aware of the existence of the functionalities.

		PC
		Frequency
Valid	1	34 (30.9%)
	2	38 (34.5%)
	3	22 (20.0%)
	4	15 (13.6%)
	5	1 (0.9%)
	Total	110 (100%)

TABLE 4.14

Table 4.14 presents the skill levels of the users regarding the Privacy Control on Facebook. This result emphasizes the urge to improve this skill.

Hypothesis H1a is accepted: Most users have low skill levels to decide to disclose personal information.

ADE KNOWLEDGE

Knowledge results: Awareness of data exploitation

The results of the *awareness of data exploitation* questions illustrate insufficient knowledge of operations of the users. Q16 and Q18 are exemplified to validate this. Q16 tested the knowledge of the users regarding the operations of advertisement options for third parties. The question specifically asked*: 'Do you think that firms that advertise on Facebook offer you personalized advertisements because they are able to download your personal information?'* Only 13.6% of the users answered this questions correctly (No, this is not possible). In other words, 86.4% of the users mistakenly think that advertises can download your information. This implicates that 86.4% of the users take into account that this occurs and yet they continue using the service. These results indicate that the users do not care. Q18 tested the knowledge of the users regarding the specific information that Facebook exploits in their data operations. The questions asked: '*Does Facebook exclusively use your profile information to offer you personalized advertisements?'* Most users (53%) think that the question is true, while only 20% correctly answers it. It is naïvely to think that a firm that sits a pile of data and exploits your data, to limit their exploitations to your profile page. Table 4.15 illustrates the ADE knowledge levels.

		ADE
		Frequency
Valid	1	58 (52.7%)
	2	37 (33.6%)
	3	10 (9.1%)
	4	5 (4.5%)
	5	0 (0.0%)
	Total	110 (100%)

TABLE 4.15

The low awareness of data exploitation score explains why the Provider Trust correlation is not discovered. The users have low levels of knowledge regarding the exploitation that

Facebook applies. Provider trust is therefore a non-issue to them. The non-critical attitude of the users make provider trust redundant.

Several participants reacted to the survey with a typical statement: 'I had the feeling that I'm not that competent in the use of Facebook' and 'I learned a few lessons by filling the survey.' These typical statements illustrate the increased awareness of the privacy risks of personal information disclosure.

The knowledge levels of the users and their personal reactions to the survey illustrate the urge to increase the knowledge regarding the awareness of data exploitation. Recommendations to achieve this are include at the end of this chapter.

Hypothesis H1b is accepted. Most users have low knowledge levels to decide to disclose personal information.

Competence scores: Perceived Benefit variables (excluded expert)

Users are better able to estimate the perceived level of enjoyment and the perceived level of self-presentation than they are able to estimate the perceived level of Relationship maintenance (Table 4.16). On average most of the users score sufficient (≥3).

Competences per variable of Perceived Benefits				
		PE	PS	PR
Competence level		Frequency	Frequency	Frequency
Valid	1	4 (3.6%)	2 (1.8%)	10 (9.1%)
	2	3 (2.7%)	4 (3.6%)	16 (14.5%)
	3	16 (14.5%)	13 (11.8%)	25 (22.7%)
	4	24 (21.8%)	24 (21.8%)	31 (28.2%)
	5	63 (57.3%)	67 (60.9%)	28 (25.5%)
	Total	110 (100%)	110 (100%)	110 (100%)

TABLE 4.16

Competence scores: Perceived Benefit variables (included expert)

The difference between the inclusion and exclusion of Expert 2 hardly makes any difference for the competences scores for *Perceived Enjoyment.* The competence in *Self-presentation* becomes more average and the *Relationship Maintenance* competences with low levels decrease (Table 4.17). The reason for this small deviation is that Expert 2 has little influence on the Predicting Factors of Benefits, since this expert is not selected for these variables.

Competence levels of variables of Perceived Benefits			
	PE	PS	PR
Competence level	Frequency	Frequency	Frequency
1	4 (3.6%)	5 (4.5%)	2 (1.8%)
2	3 (2.7%)	8 (7.3%)	7 (6.4%)
3	15 (13.6%)	15 (13.6%)	29(26.4%)
4	25 (22.7%)	38 (34.5%)	46 (41.8%)
5	63 (57.3%)	44 (40.0%)	26 (23.6%)
Total	110 (%)	110 (%)	110 (100%)

TABLE 4.17

Hypothesis H2 is accepted: Most users of Facebook have a high competence level to accurately estimate the perceived benefits of disclosing personal information.

Competences scores: Perceived Trust (included expert)

The interpretation of the competence level for the trust variable has low validity since the influence of Expert 2 on this set of variables is categorized as strong.

The competences to estimate the trust for the predicting variables are scattered over the all competence levels (Table 4.18). The competence levels 4 and 5 seem to have a slight upper hand. Since the validity of this result is insufficient no recommendations are made to improve these competence scores.

Trust per factor competence levels

		PPT	PMT	PLT	PTTP
Competence level		Frequency	Frequency	Frequency	Frequency
Valid	1	10 (9.1%)	15 (13.6%)	18 (16.4%)	17 (15.5%)
	2	20 (18.2%)	11 (10.0%)	17 (15.5%)	15 (13.6%)
	3	17 (15.5%)	27 (24.5%)	19 (17.3%)	20 (18.2%)
	4	29 (26.4%)	33 (30.0%)	26 (23.6%)	28 (25.5%)
	5	34 (30.9%)	24 (21.8%)	30 (27.3%)	30 (27.3%)
	Total	110 (100%)	110 (100%)	110 (100%)	110 (100%)

TABLE 4.18

Hypothesis H3 is rejected: Most users of Facebook have low competences in estimating the level of trust in the platform entities.

Direct analysis of main variables

Besides measuring the competences of estimation with the sub-variables, the main variables are also measured directly in the survey.

A competence scores is computed for the users for these variables. The competence scores for these variables are not computed with the predictor factors of the variables. Results of that analysis are included in the next section.

Main variable competences

		Privacy Risk	Benefits
Competence level		Frequency	Frequency
Valid	1	2 (1.8%)	15 (13.6%)
	2	3 (2.7%)	22(20.0%)
	3	22 (20.0%)	24(21.8%)
	4	18 (16.4%)	28(25.5%)
	5	65 (59.1%)	21(19.1%)
	Total	110 (100%)	110(100%)

TABLE 4.19

Table 4.19 indicates that the users have high competences in estimating the *Perceived Privacy Risks* directly (measured without the predictor variables). The estimation of the *Perceived Benefits* indicates spread results for the difference levels of competence. The results in table 4.19 are not incorporated in the overall competence scores. These are presented in the next section.

Predicted competence levels overall

In this part of the chapter the competence levels for the Perceived Privacy Risks and Perceived Benefits are predicted with the prediction models. The functions and variables to predict the main variables are included in the previous section.

The overall estimation competence scores are presented with and without Expert 2. The trust competences are disregarded since the prediction model is interpreted as insufficient. The effect of the inter-correlation between Trust and Privacy Risks (R^2=

0.052) did not significantly influenced the competence results and is therefore not separately included.

This comes down to two different results for the competence analysis:

- Competence Analysis 1; with expert exclusion (Thus, trust not included);
- Competence Analysis 2; without expert exclusion (Inter-correlation included);

The results of the Competence Analysis 2 and 3 are similar; therefore the results are presented once.

- Competence Analysis 3; without expert exclusion (Inter-correlation disregarded).

Per competence table several cells are colored red, green, yellow, grey or orange. The colors highlight different type of users. The green color illustrates users who possess high competence in the estimation of *Perceived Benefits* and *Perceived Privacy Risks.* This type of user is to a high extent able to decide whether to disclose personal information or not. Risks are known to this type of user and therefore can be avoided. This type of user is also highly competent in estimating the benefits, and therefore is competent in deciding if disclosure decisions result positively. The red color indicates the type 'hazardous users'. The users with this color are a hazard to their own privacy and privacy of others. This type of user is competent in estimating the benefits, and therefore is competent in extracting the benefits from disclosure decisions. The competence level in estimating the privacy risks is low, which is a dangerous combination. The yellow color indicates a sufficient competence level, the users in this category are labels as 'balanced users'. These users have sufficient competences on at least the privacy risk axe. This type of user is no threat to the privacy of himself or others. The grey color indicates an insufficient competence level on both axes. This type of user is assumed to disclose minimal personal information since the rewards (benefits) are not correctly recognized. When the benefit does not correspond with the executed activity, the activity is not pursued. This type of user has a low privacy risk profile. The orange color indicates the type of users that is highly competent in estimating the *Perceived Privacy Risks* but has low competence in estimating *Perceived Benefits.* This user type probably neglects the opportunity to experience benefits and therefore avoids personal information disclosure. This type of user is labeled as 'opportunity neglects'.

Competence Analysis 1

The inter-correlation between *Perceived Trust* and *Perceived Privacy Risks* is disregarded since the golden standard for the Trust variables does not meet the criteria (Appendix E).

The included predictor variables for *Perceived Benefits* are *PE* and *PR.* The included predictor variables for *Perceived Privacy Risks* are *PD* and *PR.*

Privacy Risks Comp * Benefit Comp Crosstabulation

		Benefit Competence					
		1	2	3	4	5	Total
Privacy Risk Competence	1	0	0	0	1	0	1
	2	0	0	2	4	2	8
	3	0	1	2	10	9	22
	4	2	3	10	19	21	55
	5	0	0	4	11	9	24
Total		2	4	18	45	41	110

TABLE 4.20

The results in table 4.20 indicate that most (54.55%) of the users have high levels of competences in estimating the *Perceived Benefits* and *Perceived Privacy Risks* that the users are exposed to. Only a few participants are in the hazardous area (8.18%), while 32% of the users the have sufficient competences. A minimum number of participants (4.54%) are labeled as *´opportunity neglects´*.

Competence Analysis 2

In this Competence Analysis none of the experts are excluded from the golden standard. Inter-correlation between *Perceived Trust* and *Perceived Privacy Risks* is therefore useful to include.

The included predictor variables for *Perceived Benefits* are *PE* and *PR*. The predictor variables that determine the *Perceived Privacy Risk* are *PD* and *FR*. The predictor for *Perceived Trust* (which correlates with *Perceived Privacy Risks*) is *PMT*.

The results in table 4.21 indicate that the most (57.27%) users have high competences in estimating the *Perceived Benefits* and *Perceived Privacy Risks*. The hazardous area contains 10.0% of the users, while the users that are labeled as '*opportunity neglects* are N=2. The sufficient competence area contains 30.0% of the users. One user is categorized in the grey area.

Privacy Risks Comp * Benefit Comp Crosstabulation

		Benefit Competence					
		1	2	3	4	5	Total
Privacy Risk Competence	1	0	0	0	0	2	2
	2	0	1	2	4	3	10
	3	3	1	1	4	9	18
	4	1	1	9	15	22	48
	5	0	0	6	9	17	32
Total		4	3	18	32	53	110

TABLE 4.21

Results of the previous competence analysis indicate sufficient abilities to estimate the privacy risks and benefits of personal information disclosure decisions. Approximately 9.0% of the users are labeled as ‘hazardous users’. These users solicit for privacy protection. Recommendations to protect these users are included in the next section.

4.6 Conclusion of empirical analysis

The results of three different analyses are discussed:

- The estimation competences regarding privacy risks and benefits;
- The skill levels to disclose personal information;
- The knowledge regarding data handling operations.

Tables 4.20 and 4.21 indicate that most of the users have sufficient / high competence levels regarding the estimation of the privacy risks and benefits. Tables 4.13 and 4.14 shows that most users do not possess high skill levels to use Facebook. The knowledge of most users regarding the awareness of data handling operations is insufficient (Table 4.15).

5 CONCLUSION

In this final chapter a short summary of the research results is presented. The following paragraphs tie the results together and present recommendations beyond the empirical findings of the research. At the end of the chapter the limitations of this research and the recommendations for future research are included.

5.1 Empirical findings

The validation of the conceptual model culminated in expected and unexpected results. The expected correlations between *Perceived Benefits* and the predictor variables *Perceived Enjoyment, Perceived Self-Presentation* and *Perceived Relationship Maintenance* are moderately present. Likewise, the expected predictors of *Perceived Privacy Risks* are discovered. The variables *Perceived Damage, Perceived Likelihood* and *Perceived Future Risk* moderately correlate. Analyzing the *Perceived Trust* resulted in unexpected findings. The variables *Perceived Member Trust, Perceived Legal Trust* and *Perceived Third Party Trust* are moderately present, however the *Perceived Provider Trust* is not. Even though measured on a different level of detail, the assumption of Krasnova & Veltri (2010) that this correlation would be present is not corroborated. The insufficient level of knowledge about the awareness of data exploitation of the users explains this deviation. This result has implications to the recommendations.

The findings of the competence-, skills- and knowledge analyses indicate inconsistent results. Most users possess sufficient competences in the estimation of P*erceived Benefits* and P*erceived Privacy Risks*. The small group (9%) of users who score insufficient solicits for protection. The estimation of *Perceived Trust* by the user and the experts is scattered, which indicates the complexity of quantifying the concept of trust. Other conclusions than a statement regarding the complexity of quantifying the concept of trust are not made since these would not be useful. Even though the influence of *Perceived Trust* on the privacy calculus is only slightly present, this concept is a candidate for more thorough and in depth research. The research results indicate that the privacy policy and privacy control skills of the users are insufficient. On the subject of the knowledge about the awareness of data exploitation the users score insufficient as well. This study illustrates that room for improvements by the government, organizations and the users themselves is present. Several empirical results indicate the absence of the critical attitude to assess online privacy. Although it is not the main theme of this research, the attitude of the users seems to greatly influence the disclosure behavior.

These results accumulate to the following conclusion:

- *Most users of Facebook possess high competences in estimating the privacy risks and the benefits to perform the privacy calculus, while the specific skills and knowledge to decide to disclose personal information are lacking.*

5.2 Theoretical findings

The theoretical findings of this study discuss the role of rationality in the privacy calculus. The theory of planned behavior illustrates that the willingness to execute a certain behavior is subject to the subjective norm of the individual. Social pressure applied by for instance friends and family negatively influences rationality of decision making. Kahneman & Tversky (1974) state that the decision to disclose is accomplished by the use of an automatic-pilot. Previous encounters and past experiences with the decision are recollected to determine future behavior. The automatic-pilot impairs the rationality of decision making; the privacy calculus is therefore not constantly applied in disclosure decisions. The third argument that disregards the rationality of the users to disclose personal information is brought by Herbert Simon with the Theory of bounded rationality (TBR). The TBR states that individuals are not entirely capable to make rational decisions due to the fact that they are not able to assimilate and digest all the information that is required to perform certain activities. The limited access to the complete set of information, the restricted cognitive limits and the time scope complete the arguments for bounded rationality.
The above mentioned arguments all illustrate the inadequate effect of rationality on the privacy calculus. Bounded rationality affects the disclosure decisions of internet users and therefore improving the competences, skills and knowledge is not sufficient. Protection from organizations and legislative authorities could provide the required handles to decide to disclose personal information.

5.3 Conclusion

The final conclusion includes the findings of the empirical and theoretical research. The main research question *'To which extent are Dutch adult users of Facebook able to accurately perform the privacy calculus?'* results in a limited positive verdict. Subsequently the answer to the main research question lays the foundations for the recommendations and the implications for the privacy debate.

- The Dutch adult users of Facebook have sufficient competences to perform the privacy calculus; however their knowledge and skills set regarding the privacy policy, privacy control and awareness of data exploitation are inferior. The rationality during the decision to disclose personal information is bounded which endangers the privacy of the users even more.

This conclusion is based on the sample of Dutch adult Facebook users. This result is extracted towards a more generic population (see paragraph 5.5 for the limitations). This research focused on users of SNS, while individuals are confronted with information disclosure decisions in many different situations as well. Smart watches, medical devices, smart cars, home automation and e-commerce collect more personal information than ever before. These developments are a soil for privacy infractions. This study provides evidence to assume that individuals would have difficulties to cope with these developments as well. Although disclosure on a SNS seems different compared with these trends, the same technologies and principles apply.

Individuals are exposed to unclear and complex disclosure decisions and often lack the competence, skills and knowledge to cope with these situations, more research is required to validate these assumptions. Disclosure decisions are often performed by an automatic-pilot and without the required information to accurately decide (bounded rationality). Besides the competences, skills, knowledge and attitudes improvements of the users the organizations and legislative authorities can provide enhanced protection.

5.4 Recommendations

The recommendations that adhere to this research are a result of the empirical and theoretical findings and are interpreted in a larger context than Facebook. Even though most users of the sample possess the capabilities to estimate the privacy risks and perceived benefits, a small portion (9%) does not possess these competences. This small group solicits for protection. The empirical results indicate insufficient skill levels regarding the privacy policy, privacy control and the knowledge of the awareness of data exploitation. The attitude of the users regarding the disclosure of personal information on the internet seems to be ignorant. The theoretical findings illustrate the lack of rationality in information disclosure decisions. The previously mentioned issues are all subject to improvement, in the next paragraph recommendations are presented.

Improvements can be performed by the users themselves, online information services (like Facebook) and the legislative power of governments and the European Union. Although the users are responsible for their own information disclosure behavior, arguments were given that protection beyond their personal responsibilities is required to assure their privacy.

The recommendations to improve these issues consist of several levels; an individual level, an organizational level and a legislative level. A coherent approach and collaborations between several parties are required to assure the online privacy of users in a structural manner.
On an individual level the awareness of data exploitation is the first candidate to be improved. Realization of the inabilities is for the user the first step of improvement. The users need to be informed of their competence-, skill- and knowledge levels to make them susceptible for improvements. Awareness of their inabilities uncovers the specific threats the user is exposed to. This step can be executed with a tool (see Appendix K for the prototype of the tool). When the specific inabilities are clear and the user is susceptible for improving them specific training in the Privacy Policy skill and Privacy Control skills are recommended. Preferably this training is offered independently from the particular service provider.

On a legislative level the legislative authority is recommended to oblige organizations (organizational level) to s*implify and reduce their privacy policies*. The empirical findings illustrate that the privacy policy is too complex for users to understand and apply. Symbols for instance can be used to improve the understandability. The legislative power is recommended to force organizations to apply more transparency to their data handling operations. The implementation of this recommendation should not only apply to privacy policies but also to the core features of the services. This will result in *more prominently offering privacy features* in the service.

Some of these recommendations are already suggested by the European Commission (EC) and implemented by some multinationals. These intentions are evaluated as the first step in the right direction. The EC already suggested improved legislation to assure the privacy of their citizens. Most of the suggested legislation already greatly improves the privacy and is in line with the recommendations: the *right to be forgotten* and the *right of data portability* improve the control of the users. The ability to undo certain disclosure decisions and modify / delete certain personal information improves privacy control.
The bounded rationality of the users results in endangered privacy. The EC suggests legislation that assures *privacy-by-default* and *privacy-by-design*. The privacy-by-default legislation obligates organizations to set the privacy settings of services to be secured by default. The users with bounded rationality, insufficient skills or knowledge are protected by default with this instrument. The privacy-by-design legislation obligates organizations to design their services with the privacy of the users in mind from the start of developing the service. This brings advantages that the privacy is secured from the start. Prevention of privacy infraction is the main theme in this principle. The EC tries to tackle the insufficient skill levels of the users regarding the privacy policy as well. The suggested legislation obligates organizations to inform their users as clearly, transparently and understandably as possible.
The suggested legislation also contains drawbacks. The intentions of the EC are to improve privacy protection and economic activity. The suggested legislations for instance stimulate an Export-Import-Module (EIM) by which personal information can be downloaded in a standardized manner. This standardized manner reduces switching costs but creates possible loss of data when the module is insufficiently protected. Another point for discussion is the *Provider Trust*. In this research the provider trust was not discovered, while the EC fundaments their arguments (Why do we need an EU data protection reform, 2012) to improve economic activity with the presence of this component. Extra research is recommended to clear this distinction.

Some organizations lead the way by already implementing multiple features by which they satisfy the upcoming legislation. Google implemented the *'right to be forgotten' form* and Facebook constantly reforms their privacy features like control settings and the privacy policy. A mild version of privacy-by-default is already implemented. The latest update (November 2014) by Facebook deleted approximately 70% of the privacy policy. The policy is also color coded to reduce complexity and increase transparently and understandability. This is a viable alternative solution to the symbols in the privacy statements. Facebook also implemented an alternative EIM which does not match the intentions of the EC. Most changes are in line with the suggested improvements by the EC. Facebook anticipates on the upcoming legislation by implementing these features, but infracts some others by changing the core content of the policy. The social network will share personal data with other services (like Whatsapp) on a larger scale.

Implications to the privacy debate are that the intentions of the EC are valuable additions to the privacy protection of individuals but paradoxically also could fundament more infractions on the long run. The intentions of the organizations to enhance privacy control are admirable but need to be monitored continuously to prevent infractions. Altogether, the individuals require to be armed with privacy protection continuously

since the information environment rapidly evolves. The protection by the EC and individual organizations are the first step in the right direction, but continuously need to be improved and monitored.

5.5 Limitations

In this paragraph two limitations that slightly influence this research are mentioned. Firstly, in this research the Dutch adult Facebook users are the unit of observation. In the recommendation part of this report recommendations regarding a larger population are given. The sample results attempt to be extracted to all Facebook users and later on even to all individuals that disclose information on the internet. Generic statements can be made about these groups however specific recommendations are not possible for two reasons: a) the sample slightly deviates from the entire population and b) the sample does not contain children and elderly. These two elements need to be taken in account when interpreting the results.

Secondly, the competence-, skill- and knowledge levels of the users are categorized by calculation the deviations from the golden standard. The golden standard is structured with the objective interpretations of the experts. Multiple arguments for bounded rationality are given. Bounded rationality affects the experts and the users and thereby their scorings as well. This limitation is minimalized by the use of specific criteria (Appendix E).

5.6 Future research

Signals were given regarding the length, duration and depth difficulties of this research. This research tries to cover a broad spectrum of topics while it was limited by these factors. A recommendation for future research therefore is to split up the research into smaller parts. These smaller parts can be more thoroughly investigated. In this research the variables are often researched with 3 to 5 questions, which only covers the main essence of the variable. Since the research is executed on certain level of detail (disclosure on Facebook) the conclusions are hard to generalize towards a larger population. A detailed research per variable is proposed to explain the relationships and variables even more.

The second recommendation for future research aims to broaden the research sample. The research sample in this investigation consists of Dutch adult Facebook users. Children and adolescents are deliberately excluded. However, these users groups of Facebook are a major part of the user base of the organization. Especially the younger users of Facebook (and other social media) are said to possess low competences to safely disclose personal information. It is stated that these users perform significantly more risky behavior. Extended research with these users would result in more effective knowledge to protect them in the future.

6 ACKNOWLEDGEMENTS

I would like to express my very great appreciation to my supervisors Fons Wijnhoven and Lesley Broos. The time invested by Fons was remarkably great. Both supervisors participated in brainstorm session regarding the topics of this research. I believe this is an important moment in time for the (privacy) information management discussion. I am glad that I could participate in the body of knowledge of it. Fons brought many ideas and ambitions to the table which sparked me to perhaps pursue a career in research.

I would like to offer my special thanks to the members of the expert panel. I am grateful that they took the time to invest in my research; on the other hand I understand that other experts declined the request to participate.
Subsequently I would like to acknowledge the effort that the participants of the survey took. The time invested by my friends, family and acquaintances was extensive.

REFERENCES

Acquisti, A., & Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Privacy enhancing technologies* (pp. 36–58). Springer. Retrieved from http://link.springer.com/chapter/10.1007/11957454_3

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, *52*(1), 27–58.

Ajzen, I. (2002). Perceived Behavioral Control, Self-Efficiacy, Locus of Control, and the Theory of Planned Behavior. *Journal of Applied Social Psychology*, *2002*(32), 665 – 683.

Ajzen, I. (2003). Theory of Planned Behavior. *Social Psychologyvolume Ⅰ: Social Cognition and Social Perception*, 347–377.

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health*, *26*(9), 1113–1127. doi:10.1080/08870446.2011.613995

Ajzen, I., & Fishbein, M. (1970). The Prediction of Behavior from Attitudinal and Normative Variables. *Journal of Experimental Social Psychology*, *1970*(6), 466 – 487.

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, *22*(5), 453–474.

Awad, N. F., & Krishnan, M. S. (2006). The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to be Profiled Online for Personalization. *MIS Quarterly*, *30*(1). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=02767783&AN=19754858&h=5b3rQaexq4CziktvxAfIe%2BeVpRb81BhlgYaHugsmyZrmg4cuJfuuGmwOyu2w%2B73%2BWrpIoMW1X6iMlmqwbP3%2Frg%3D%3D&crl=c

Beldad, A., de Jong, M., & Steehouder, M. (2010). Reading the least read? Indicators of users' intention to consult privacy statements on municipal websites. *Government Information Quarterly*, *27*(3), 238–244. doi:10.1016/j.giq.2010.01.004

Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, *27*(3), 264–271. doi:10.1016/j.giq.2010.03.001

Bonsón, E., Torres, L., Royo, S., & Flores, F. (2012). Local e-government 2.0: Social media and corporate transparency in municipalities. *Government Information Quarterly*, *29*(2), 123–132. doi:10.1016/j.giq.2011.10.001

Boyd, D. (2007). Why youth (heart) social network sites: The role of networked publics in teenage social life. *MacArthur Foundation Series on Digital learning–Youth, Identity, and Digital Media Volume*, 119–142.

boyd, d. (2008). Facebook's Privacy Trainwreck: Exposure, Invasion, and Social Convergence. *Convergence: The International Journal of Research into New Media Technologies*, *14*(1), 13–20. doi:10.1177/1354856507084416

boyd, danah m., & Ellison, N. B. (2007). 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

Che, D., Safran, M., & Peng, Z. (2013). From Big Data to Big Data Mining: Challenges, issues, and Opportunities. In *Database Systems for Advanced Applications* (pp. 1–15). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-40270-8_1

Culnan, M. J. (1993). " How Did They Get My Name?": An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use. *Mis Quarterly*, *17*(3). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=02767783&AN=9404050726&h=%2FAU2WYZFq067qAEcPyZVf3JsLiohXMYdRMvp866u9J%2FcY304nYG0QPxu%2FCJH8oaRYSoT6B5mISHNgZi1PFDO1g%3D%3D&crl=c

Culnan, M. J., & Armstrong, P. K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science*, *1999*(Vol 10, No, 1), 104 – 115.

Culnan, M. J., & Bies, R. J. (2003). Consumer privacy: Balancing economic and justice considerations. *Journal of Social Issues*, *59*(2), 323–342.

Davison, H. K., Maraist, C., & Bing, M. N. (2011). Friend or Foe? The Promise and Pitfalls of Using Social Networking Sites for HR Decisions. *Journal of Business and Psychology*, *26*(2), 153–159. doi:10.1007/s10869-011-9215-8

Debatin, B., Lovejoy, J. P., Horn, A.-K., & Hughes, B. N. (2009a). 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

Debatin, B., Lovejoy, J. P., Horn, A.-K., & Hughes, B. N. (2009b). 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. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research*, *17*(1), 61–80. doi:10.1287/isre.1060.0080

Dubbink, W., Graafland, J., & Liedekerke, L. (2008). CSR, Transparency and the Role of Intermediate Organisations. *Journal of Business Ethics*, *82*(2), 391–406. doi:10.1007/s10551-008-9893-y

Eling, N., Krasnova, H., Widjaja, T., & Buxmann, P. (2013). Will you accept an APP? Empirical investigation of the decisional calculus behind the adoption of application on Facebook. *Security and Privacy of Information and IS*, 1 – 30.

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

Facebook cijfers juli 2013, 2013, Retrieved August 13, 2014, from http://www.likeconomics.nl/2013/07/facebook-cijfers-groei-in-alle-leeftijden-en-82-miljoen-accounts/

Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, *9*(2), 319–342. doi:10.1177/1461444807075015

Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, *25*(1), 153–160. doi:10.1016/j.chb.2008.08.006

Fuchs, C. (2012). The Political Economy of Privacy on Facebook. *Television & New Media*, *13*(2), 139–159. doi:10.1177/1527476411415699

Gauch, S., Speretta, M., Chandramouli, A., & Alessandro, M. (2007). User Profiles for Personalized Information Access. *The Adaptive Web*, *2007*, 54 – 89.

Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007a). User profiles for Personalized Informaation Acces. In *The Adaptive Web Methods and Strategies of Web Personalization* (2007th ed., pp. 54 – 89). Springer- Verlag Berlin Heidelberg.

Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007b). User profiles for personalized information access. In *The adaptive web* (pp. 54–89). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-72079-9_2

Geva, T., Oestricher-Singer, G., Efron, N., & Shimshoni, Y. (n.d.). Do customers speak their minds. Using forums and search for prediction sales. *Knowledge Management and Business Intelligence*, *2013*, 1 – 17.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, *103*(4), 650.

Groot, R. E. J. (2012). IT-based risks in advergame campaigns: a focus on fairness and privacy. Retrieved from http://essay.utwente.nl/61956/

Hill, S., Benton, A., & Van den Bulte, C. (2013). When Does Social Network-Based Prediction Work? A Large Scale Analysis of Brand and TV Audience Engagement by Twitter Users. *Knowledge Management and Business Intelligence*, *2013*, 1 – 16.

Hinduja, S., & Patchin, J. W. (2008). Personal information of adolescents on the Internet: A quantitative content analysis of MySpace. *Journal of Adolescence*, *31*(1), 125–146. doi:10.1016/j.adolescence.2007.05.004

Horst, M., Kuttschreuter, M., & Gutteling, J. M. (2007). Perceived usefulness, personal experiences, risk perception and trust as determinants of adoption of e-government services in The Netherlands. *Computers in Human Behavior*, *23*(4), 1838–1852. doi:10.1016/j.chb.2005.11.003

IFIP WG 9.2, 9.6/11.6, 11.7/FIDIS International Summer School on the Future of Identity in the Information Society, Fischer-Hübner, Simone. (2011). *Privacy and identity management for life: 6th IFIP WG 9.2, 9.6/11.7, 11.4, 11.6/PrimeLife International Summer School, Helsingborg, Sweden, August 2-6, 2010, revised selected papers*. London; New York: Springer.

Jabr, W., & Rahman, M. (2013). The Market for Information: An Analysis of the Online Word-of-Mouth, 1 – 16.

Jaeger, P. T., & Bertot, J. C. (2010). Transparency and technological change: Ensuring equal and sustained public access to government information. *Government Information Quarterly*, *27*(4), 371–376. doi:10.1016/j.giq.2010.05.003

Johnson, T. J., & Kaye, B. K. (2010). Believing the blogs of war? How blog users compare on credibility and characteristics in 2003 and 2007. *Media, War & Conflict*, *3*(3), 315–333. doi:10.1177/1750635210376591

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. doi:10.1037/0003-066X.58.9.697

Kehr, F., Wentzel, D., & Mayer, P. (2013). Rethinking the Privacy Calculus: On the Role of Dispositional Factors and Affect. Retrieved from http://aisel.aisnet.org/icis2013/proceedings/ResearchInProgress/56/

Keith, M. J., Thompson, S. C., Hale, J., Lowry, P. B., & Greer, C. (2013). Information disclosure on mobile devices: Re-examining privacy calculus with actual user behavior. *International Journal of Human-Computer Studies*, *71*(12), 1163–1173. doi:10.1016/j.ijhcs.2013.08.016

Knijnenburg, B. P., Kobsa, A., & Jin, H. (2013). Counteracting the negative effect of form auto-completion on the privacy calculus. *Submitted to ICIS*. Retrieved from http://www.ics.uci.edu/~kobsa/papers/2013-ICIS-Kobsa.pdf

Krasnova, H., & Veltri, N. F. (2010). Privacy calculus on social networking sites: Explorative evidence from Germany and USA. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (pp. 1–10). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5428447

Krasnova, H., Veltri, N. F., & Günther, O. (2012). Self-disclosure and Privacy Calculus on Social Networking Sites: The Role of Culture: Intercultural Dynamics of Privacy Calculus. *Business & Information Systems Engineering*, *4*(3), 127–135. doi:10.1007/s12599-012-0216-6

Lanier, J (2013), Who owns the future?

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. doi:10.1518/hfes.46.1.50_30392

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

Livingstone, S. (2008a). 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. (2008b). 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

Ljepava, N., Orr, R. R., Locke, S., & Ross, C. (2013). Personality and social characteristics of Facebook non-users and frequent users. *Computers in Human Behavior*, *29*(4), 1602–1607. doi:10.1016/j.chb.2013.01.026

Lwin, M. O., & Williams, J. D. (2003). A model integrating the multidimensional developmental theory of privacy and theory of planned behavior to examine fabrication of information online. *Marketing Letters*, *14*(4), 257–272.

MacDonald, J., Sohn, S., & Ellis, P. (2010). Privacy, professionalism and Facebook: a dilemma for young doctors: Social networking and professionalism. *Medical Education*, *44*(8), 805–813. doi:10.1111/j.1365-2923.2010.03720.x

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

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

McIvor, R., McHugh, M., & Cadden, C. (2002). Internet technologies: supporting transparency in the public sector. *International Journal of Public Sector Management*, *15*(3), 170–187. doi:10.1108/09513550210423352

Miyazaki, A. D., & Fernandez, A. (2000). Internet privacy and security: An examination of online retailer disclosures. *Journal of Public Policy & Marketing*, *19*(1), 54–61.

Muise, A., Christofides, E., & Desmarais, S. (2009). More information than you ever wanted: Does Facebook bring out the green-eyed monster of jealousy? *CyberPsychology & Behavior*, *12*(4), 441–444.

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, *41*(1), 100–126.

Olivero, N., & Lunt, P. (2004). Privacy versus willingness to disclose in e-commerce exchanges: The effect of risk awareness on the relative role of trust and control. *Journal of Economic Psychology*, *25*(2), 243–262. doi:10.1016/S0167-4870(02)00172-1

Palen, L., & Dourish, P. (2003). Unpacking privacy for a networked world. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 129–136). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=642635

Pariser, E. (2012), The Filter bubble: What the internet is hiding from you.

Payne, R. W., Committee, L. A. T., & others. (2003). *The guide to GenStat release 7.1*. Lawes Agricultural Trust.

Sheldon, P. (2008). The Relationship Between Unwillingness-to-Communicate and Students' Facebook Use. *Journal of Media Psychology: Theories, Methods, and Applications*, *20*(2), 67–75. doi:10.1027/1864-1105.20.2.67

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 99–118.

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information Privacy: Measuring Individuals' Concerns About Organizational Practices. *MIS Quarterly*, *20*(2). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=02767783&AN=9610124512&h=ifkwim3MagbiPw%2B9q6h4fDRgcfSLdQAkCJTIIvqtteHpp9WyUwLi8f8OIoUrFd7y9YXm2QlkeQ9vGW8W7Iz2dg%3D%3D&crl=c

Social Networking Fact Sheet, Retrieved August 13, 2014, from http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/

Social Networking sites and our lives, 2011, Retrieved August 13, 2014, from http://www.pewinternet.org/2011/06/16/social-networking-sites-and-our-lives/

Special, W. P., & Li-Barber, K. T. (2012). Self-disclosure and student satisfaction with Facebook. *Computers in Human Behavior*, *28*(2), 624–630. doi:10.1016/j.chb.2011.11.008

Sundar, S. S., & Nass, C. (2000). Source Orientation in Human-Computer Interaction: Programmer, Networker, or Independent Social Actor. *Communication Research*, *27*(6), 683–703. doi:10.1177/009365000027006001

The Role of Push–Pull Technology in Privacy Calculus.pdf. (n.d.).

Torres, L., Pina, V., & Acerete, B. (2006). E-Governance Developments in European Union Cities: Reshaping Government's Relationship with Citizens. *Governance*, *19*(2), 277–302.

Vaccaro, A., & Madsen, P. (2009). Corporate dynamic transparency: the new ICT-driven ethics? *Ethics and Information Technology*, *11*(2), 113–122. doi:10.1007/s10676-009-9190-1

Van Eerde, W., & Thierry, H. (1996). Vroom's expectancy models and work-related criteria: A meta-analysis. *Journal of Applied Psychology*, *81*(5), 575.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=02767783&AN=10758835&h=mliq1P3bTOmb3tIqmfWrYwsqorLnoPj5

dmzLIRZRztYtYLexOc%2FaX8z0KTljRTvbhS8WzOrtF7UwWyjwkXg3Qg%3D%3D&crl=c

Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science*, *7*(3), 203–220. doi:10.1177/1745691612442904

Why do we need an EU data protection reform fact sheet (2012), Retrieved from (http://ec.europa.eu/justice/data-protection/document/review2012/factsheets/1_en.pdf).

Xu, H., Luo, X. (Robert), Carroll, J. M., & Rosson, M. B. (2011). The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision Support Systems*, *51*(1), 42–52. doi:10.1016/j.dss.2010.11.017

Zhu, K. (2004). Information Transparency of Business-to-Business Electronic Markets: A Game-Theoretic Analysis. *Management Science*, *50*(5), 670–685. doi:10.1287/mnsc.1040.0226

APPENDIX A

Competence questions

Privacy statements results

The privacy policy table indicates the number of correct answers for the privacy policy questions. None of the participants answers all 5 questions correct.

Privacy policy

	Frequency	Percent	Valid Percent	Cumulative Percent
0	23	20,9	20,9	20,9
1	44	40,0	40,0	60,9
2	25	22,7	22,7	83,6
3	13	11,8	11,8	95,5
4	5	4,5	4,5	100,0
Total	110	100,0	100,0	

Privacy control on Facebook

In the privacy control table the results of the questions of the questions regarding the privacy control on Facebook are presented.

Privacy Control

Frequency of correct answers

	Frequency	Percent	Valid Percent	Cumulative Percent
0	10	9,1	9,1	9,1
1	24	21,8	21,8	30,9
2	38	34,5	34,5	65,5
3	22	20,0	20,0	85,5
4	15	13,6	13,6	99,1
5	1	,9	,9	100,0
Total	110	100,0	100,0	

Awareness of Data exploitation

Awareness of data exploitation

	Frequency	Percent	Valid Percent	Cumulative Percent
0	17	15,5	15,5	15,5
1	41	37,3	37,3	52,7
2	37	33,6	33,6	86,4
3	10	9,1	9,1	95,5
4	5	4,5	4,5	100,0
Total	110	100,0	100,0	

APPENDIX B

Data trimming criteria

Four screening criteria rules used to trim the data set:

1. **Respondents who are younger than 18 are deleted.** Children and adolescents are not part of the aimed research sample. Question 1 asks for the age group the respondent is in. Five respondents filled out 'Younger than 18', therefore n=5 are deleted.
2. **Respondents who never visit Facebook (website or application) are deleted.** The survey questions are adjusted to the use of Facebook. Respondents who do not visit the website or the application do not match the required criteria. Individuals who do not visit the Facebook website of application are not assumed are users of Facebook. Therefore these individuals are not able to answer certain questions correctly. In the data set two cases were recorded which do not match this condition, n=2 are deleted.
3. **Cases with missing data are deleted.** No cases with missing data are recorded since the Qualtrics tool forced participants to answer the questions.
4. **Outlier are not analyzed (case is not deleted).** Outliers are data points which extremely vary from the total sample. Outliers are the result of experimental error or variability in the measurement. Users that did not filled out the survey seriously or did not understand the question could be outliers. Multiple respondents who did not understand the meaning of the question are labeled as outliers. No indication of such group was found. Outlier scores from the survey are discarded for statistical analysis.

   The outlier analysis is performed by using SPSS. The scores of all questions of all cases are transformed into Z-scores. The Z-scores indicate the deviation from the mean per question. The minimum and maximum values (z-scores) of each question are analyzed. Minimum and maximum scores above and below Z=3.29 indicate outliers. Z-scores in outside the range of +3.29 or -3.29 indicate values larger or smaller than 0.01% of the sample.

Question	Score	Action
Q_13_2	N=1 scored LIKERT scale 1	Data point between 0..1 are not used for analysis.
Q55	N=2 scored LIKERT scale 2, all others scored 7.	Data point between 0..2 are not used for analysis.
Q57	N=1 scored LIKERT scale 1, all others scored higher.	Data point between 0..1 are not used for analysis.
Q78_3	N=1 scored LIKERT scale 2, all others scored higher.	Data point between 0..2 are not used for analysis.
Q80_4	N=1 scored LIKERT scale 1, all others scored higher.	Data point between 0..1 are not used for analysis.
Q82_4	N=1 scored LIKERT scale 1, all others scored higher.	Data point between 0..1 are not used for analysis.
Q84_3	N=1 scored LIKERT scale 1, all others scored higher.	Data point between 0..1 are not used for analysis.

APPENDIX C

Factor analysis procedure

The KMO statistics indicates the proportion of variance in the variables that might be caused by underlying factors. The KMO statistics is a value between 0 and 1. A value larger than 0,6 is an indicator for a probable significant component. The value of 0,6 KMO is the first argument to proceed the analysis to derive components. The Initial Eigenvalues illustrate the percentage of variance that is explained by the factor in a possible component. An Eigenvalue of larger than 1.0 is used as the second indicator to determine the number of components. Subsequently the number of factors with a larger Eigenvalue of 1,0 are validated, these explain most of variance. If most of variance is explained by the factors the influence of the factors is sufficient. The Screenplot analysis is used to estimate the number of components via a plot based on the Eigenvalue table. The screeplot illustrates the variance explained, each possible component is presented as a breaking point in the graph. This is used as the fourth argument to determine the number of components. Next is the interpretation of the Component Matrix and the Pattern Matrix. The component matrix is checked for factor coefficients larger than 0,4 and the Pattern Matrix is checked for the number of regression coefficients (components with 3 or more loadings are indicated as possible components).

The steps above are used to determine the number of components that the factors show. The analysis is run a second time. This time the determined number of components is adjusted in the analysis to a fixed number. This step provides the interpretation of the Component Correlation Matrix to confirm the chosen number of components. If the correlation between the chosen number of components is larger than 0,3 the strength of the relationship is interpreted as strong. A low strength of correlations among components indicate similar solutions of the extractions.

The final phase of the PCA is increasing the efficiency of measurement and improving the strength of the model. This phase is initiated by comparing the effects of factors on components. Negative factor – component correlations can be discarded to improve the strength of the model. The Communality table illustrates the variance that is explained by each factor in each component. Values less than 0,3 are categorized as low. The categorized items are subject to comparison with the lowest loadings from the Pattern Matrix. Items that suffice the criteria of low explaining variance values and low factor loadings are deleted to check if the model improves. Basically this comes down to creating a component without this factor or set of items and do a Linear Regression Analysis with the overlaying variable. To validate the model the correlations before and after the PCA are compared. When the regression between the factor (component) and the variable increases (compared to the earlier executed correlation analysis) the model improves.

Execution re-runs

In the paragraph on page 24 the digression about Factor Analysis the number of Factor analysis is stated at five. The number of factor analysis explained in the previous paragraph about procedure indicates only three Factor Analysis. The discrepancy between these numbers is explained by the required re-run of PCA for the factors of the variable Perceived Privacy Risk. The factors of this variable did not indicate any overlaying components. The measurement items regarding Privacy Damage, Privacy Likelihood and Future Risk are combined for analysis and the knowledge questions about Privacy Statements, Awareness of Data Exploitation and Privacy Control are combined.

Linear regression analysis

Linear Regression Analysis is used to validate the proposed model and possible components. The assumed predictive factors are tested for correlations with the main variables. This step in the procedure results in the exclusion of factors, which improves the strength of the model. This validation step is executed twice. Firstly the model as suggested is tested, secondly the model with the improvements by the PCA is tested. The model with the highest goodness of fit of regression is set as the model to proceed with. This model is used for the competence analysis part of this research.

APPENDIX D

Survey overview

Part	Subject / variable	Questions	Type of questions	
0	Demographic characteristics	First 9 questions (no number assigned)	Multiple choice / LIKERT SCALE /	
1	*Perceived damage* and *perceived likelihood*	1 till 5	LIKERT SCALE	Textual hypothetical situations
2	*Perceived future Risk*	6 till 10	LIKERT SCALE	Textual hypothetical situations
3	*Privacy policy*	11 till 15	True / False	Statements
4	*Awareness of exploitation*	16 till 20	True / False	Statements
5	*Privacy control*	21 till 25	True / False	Statements
6	*Perceived privacy risk*	26 till 30	LIKERT SCALE	Hypothetical images
7	*Trust in entities*	31 (with 12 statements)	LIKERT SCALE	Statements
8	*Perceived trust*	32 till 34	LIKERT SCALE	Hypothetical images
9	*Perceived benefits:* - *Perceived benefits* - *Enjoyment* - *Self-presentation* - *Relationship maintenance*	35 till 38	LIKERT SCALE	Hypothetical images

APPENDIX E

Panel scores

Panel scores with Expert 2 excluded

Descriptive Statistics

	N	Min	Max	Mean	Std. Deviation
Perceived Damage	4	2,00	5,60	4,1000	1,51877
Perceived Likelihood	5	5,20	7,00	5,8400	,71274
Perceived Future Risk	4	3,60	5,60	4,9500	,91469
Perceived Privacy Risk	3	4,60	5,60	5,0000	,52915
Perceived Provider Trust	**1***	5,33	5,33	5,3333	.
Perceived Member Trust	**1***	4,33	4,33	4,3333	.
Perceived Legal Trust	**1***	6,00	6,00	6,0000	.
Perceived Third Party Trust	**1***	3,00	3,00	3,0000	.
Perceived Benefits	2	4,80	5,60	5,2000	,56569
Perceived Enjoyment	2	5,60	5,80	5,7000	,14142
Perceived Self-presentation	2	5,60	6,00	5,8000	,28284
Perceived Relationship Maintenance	2	5,20	6,20	5,7000	,70711

*Score does not match criteria.

Requirements used for expert selection:

Per variable the minimum number of experts is 2;

The panel score standard deviation is smaller than 1,5 (preferably smaller than 1);

The minimum and maximum score preferably do not vary more than 2.

Panel scores with Expert 2 excluded

Descriptive Statistics

	N	Min	Max	Mean	Std. Deviation
Perceived Damage	5	2,00	7,00	4,6800	**1,84716***
Perceived Likelihood	6	5,20	7,00	6,0333	,79415
Perceived Future Risk	5	3,60	6,80	5,3200	1,14543
Perceived Privacy Risk	4	4,60	6,80	5,4500	,99833
Perceived Provider Trust	2	1,00	5,33	3,1667	**3,06413***
Perceived Member Trust	2	1,00	4,33	2,6667	**2,35702***
Perceived Legal Trust	2	1,00	6,00	3,5000	**3,53553***
Perceived Third Party Trust	2	1,00	3,00	2,0000	1,41421
Perceived Benefits	3	1,00	5,60	3,8000	**2,45764***
Perceived Enjoyment	3	5,60	6,00	5,8000	,20000
Perceived Self-presentation	3	5,60	6,60	6,0667	,50332
Perceived Relationship Maintenance	3	5,20	6,60	6,0000	,72111

APPENDIX F

Survey

Bedankt dat u deelneemt aan dit onderzoek. Dit onderzoek gaat over de persoonlijke informatie die internetgebruikers op het internet plaatsen in combinatie met privacy. Het onderzoek bestaat uit 10 korte onderdelen. Elk onderdeel bevat meestal 5 vragen. U wordt per onderdeel gevraagd om verschillende variabelen in te schatten. De meeste vragen betreffen uw gebruik van Facebook.

Alle antwoorden die u geeft worden uitsluitend gebruikt ten gunste van de wetenschap. Dit onderzoek wordt anoniem afgenomen. Er wordt niet om uw naam gevraagd.

Het onderzoek duurt ongeveer 15 minuten. U kunt nu starten met het eerste onderdeel.

Hoe oud bent u?	Jonger dan 18 18 – 25 26 – 35 36 – 45 46 – 55 56 – 65 Ouder dan 65
Wat is uw geslacht?	Man Vrouw
Wat is uw hoogst afgeronde opleiding (of waar bent u momenteel in ingeschreven)?	Geen Lagere school / Basis Onderwijs Lager beroepsonderwijs of voorbereidend middelbaar beroepsonderwijs (LBO of VMBO) Middelbaar beroepsonderwijs (MBO) Hoger algemeen voortgezet onderwijs of voorbereidend wetenschappelijk onderwijs (HAVO of VWO) Hoger beroepsonderwijs (HBO) Wetenschappelijk onderwijs (WO)
Hoe vaak bezoekt u de website of applicatie (app) van Facebook voor privégebruik?	Nooit 1x per maand 1x per week Meerdere keren per week 1x per dag Meerdere keren per dag
Hoe vaak deelt u iets op Facebook?	Nooit 1x per maand 1x per week Meerdere keren per week 1x per dag Meerdere keren per dag
Wat is de voornaamste reden dat u Facebook voor privégebruik gebruikt?	Ik gebruik Facebook niet Om op de hoogte te blijven Omdat iedereen dat doet Contact met vrienden en familie Om spelletjes te spelen Tegen de verveling Om anderen te informeren over mijzelf Anders, namelijk..
In welke mate schat u uw kennis en kunde betreffende het gebruik en misbruik van uw persoonlijke informatie op Facebook?	Geen kennis en kunde 1-2-3-4-5-6-7 Volledige kennis en kunde

In welke mate vindt u online privacy belangrijk?	Niet belangrijk 1-2-3-4-5-6-7 Erg belangrijk
Heeft u de privacyverklaring / privacyovereenkomst van Facebook gelezen?	Ja, volledig Ja, gedeeltelijk Nee

De volgende vragen gaan over het inschatten van de privacyschade en de waarschijnlijkheid van bepaalde situaties. **Een aantal situaties komen dagelijks voor en andere situaties zijn verzonnen**. Geef per situatie aan in welke mate jij de schade schat en in welke mate jij de waarschijnlijkheid van het daadwerkelijk gebeuren van die situatie schat. Er zijn geen goede of foute antwoorden. Dit onderdeel bestaat uit 5 vragen.

Betekenis / definitie:
Privacyschade: de financiële-, reputatie- of psychologische schade die jij ervaart door de situatie.
Waarschijnlijkheid: de aannemelijkheid dat de beschreven situatie daadwerkelijkheid gebeurt.

1. Jouw profielinformatie van Facebook wordt commercieel gebruikt door adverteerders. Een adverteerder biedt jou een gepersonaliseerde advertentie aan onder andere op basis van jouw leeftijd, geslacht, status van je relatie, hobbies en de pagina's die je hebt geliked.

2. Gebruikers van Facebook die jij niet kent (bijvoorbeeld vrienden van vrienden of andere onbekenden) zien foto's die op Facebook zijn geüpload waar jij op staat afgebeeld. Deze foto's kunnen zijn geüpload door jou of door anderen.

3. De overheid verzamelt gegevens over jou en heeft een nauwkeurig profiel van jou opgebouwd via onder andere jouw Facebook profiel pagina en Facebook-gebruik. Op deze manier bewaakt de overheid de maatschappelijke veiligheid.

4. Over het algemeen stuurt Facebook jouw persoonlijke informatie door naar adverterende bedrijven. Met deze informatie kun je direct persoonlijk geïdentificeerd worden. Het betreft hier gegevens als bijvoorbeeld jouw naam en email-adres.

5. Facebook heeft inzicht in de websites die jij hebt bezocht en links die jij hebt aangeklikt. Deze informatie wordt gebruikt voor commerciële doeleinden.

The image below illustrates the scales used.

In dit onderdeel van de vragenlijst worden een aantal situaties gepresenteerd. **Een aantal situaties zijn daadwerkelijk gebeurd en een aantal situaties zijn verzonnen.** Lees de situaties zorgvuldig en beoordeel de mate van de toekomstige waarschijnlijkheid. Dit onderdeel bestaat ook uit 5 vragen.

Betekenis / definitie:
De **toekomstige waarschijnlijkheid** is: de kans dat deze situatie in de toekomst jou overkomt.

6. Stel je de volgende situatie voor:

Een kledingwinkel waar jij graag kleren koopt, bepaalt op basis van de kleding die jij in het verleden hebt gekocht en de baan die jij uitvoert jouw financiële vermogen. Het bedrijf toont bepaalde producten of aanbiedingen niet meer aan jou omdat uit jouw persoonlijke informatie is gebleken dat dit niet past bij jouw financiële vermogen (m.a.w. je hebt te veel of te weinig geld).

7. Stel je de volgende situatie voor:

Je laat je op Facebook negatief uit over de één-kind-politiek in China. In China was het lang niet toegestaan om meer dan één kind te hebben. Je zet op Facebook dat je het een stomme regel vindt. Een paar jaren later ga je op vakantie naar China. Op het vliegveld word je geweigerd vanwege je eerdere uitspraak op Facebook betreffende de één-kind-politiek.

8. Stel je de volgende situatie voor:

Je laat je negatief uit over het bedrijf waar je werkt. Je zet op Facebook dat het bedrijf slecht omgaat met haar werknemers. Op het moment dat je aan het solliciteren bent voor een nieuwe baan krijgt jouw potentieel toekomstige werkgever inzicht in jouw negatieve uitlatingen betreft jouw vorige bedrijf. Om deze reden nodigt de potentiële werkgever jou niet uit voor een kennismakingsgesprek. Je loopt hierdoor een potentiële baan mis.

9. Stel je de volgende situatie voor:

Op Facebook heb jij foto's staan van jouw sportactiviteiten en ook informatie over hoe gezond of ongezond jij leeft. Denk hierbij aan foto's waarop jij alcohol consumeert of informatie betreffende hoe gezond jouw dieet is. Jouw verzekeringsmaatschappij gaat onder andere uit deze foto's en informatie afleiden hoe gezond jij leeft. De verzekeringsmaatschappij gebruikt mede deze informatie om de

prijs van jouw premiebedrag van jouw verzekering te gaan bepalen. Gezonde mensen gaan minder betalen en ongezonde mensen gaan meer betalen.

10. Stel je de volgende situatie voor:

Jij gebruikt Facebook al 5 jaar. Je hebt al veel informatie gedeeld en veel foto's geüpload. Een onbekende hacker steelt jouw afgeschermde persoonlijke informatie van Facebook en gebruikt deze voor kwaadaardige doeleinden zonder dat jij hier van af weet.

The image below illustrates the scale present with each question:

Dit onderdeel van de vragenlijst gaat over privacyverklaringen. Er zijn een aantal situaties gegeven. **Een aantal zijn op waarheid berust en een aantal zijn verzonnen of aangepast.** Zijn de volgende uitspraken Waar of Niet waar? Als je het antwoord niet denkt te weten, gok dan niet. Vul dan Weet ik niet in. Dit onderdeel bestaat uit 5 vragen.

11. Facebook is de eigenaar van het IE-recht (recht op intellectuele eigendom zoals geschreven teksten en foto's) betreffende de informatie, updates, foto's en video's die jij op Facebook deelt.

12. Lees de volgende tekst en beoordeel of deze in zijn volledigheid (of met dezelfde betekenis) in de privacyverklaring van Facebook staat:

13. De volgende tekst komt letterlijk uit de privacyovereenkomst van Facebook. Lees de tekst en beantwoord de vraag.

"Je geeft ons toestemming je naam, profielfoto, inhoud en informatie (zoals een merk dat je leuk vindt) te gebruiken voor commerciële, gesponsorde of gerelateerde inhoud die door ons wordt aangeboden of verbeterd. Dit betekent bijvoorbeeld dat je een bedrijf of een andere partij toestaat ons te betalen om je naam en/of profielfoto met je inhoud of informatie te tonen, zonder dat je daarvoor compensatie krijgt."

Stel je voor dat de volgende situatie zich voordoet:
Facebook stuurt jouw profielfoto door naar een advertentiebedrijf die vervolgens met jouw profielfoto advertenties maakt en deze aanbiedt aan jouw vrienden. Beoordeel de volgende uitspraak:

Facebook mag dit doen volgens haar eigen privacyverklaring.

14. Staat de volgende verklaring, of een verklaring met dezelfde betekenis in de Facebook privacyovereenkomst?

"We geven onze advertentiepartners of klanten pas gegevens nadat we je naam en andere elementen die je persoonlijk zouden kunnen identificeren hebben verwijderd, of wanneer we de gegevens zodanig hebben gecombineerd met de gegevens van anderen dat ze jou niet meer persoonlijk identificeren."

Beoordeel de volgende bewering:
De bovenstaande verklaring, of een verklaring met dezelfde betekenis staat in de Facebook privacyovereenkomst.

15. Stel je voor dat de onderstaande tekst uit de privacyverklaring van Facebook komt. Lees de tekst en beoordeel de onderstaande bewering.

"Je blijft de eigenaar van al je gegevens, zelfs al geef je ons toestemming de gegevens die we van je hebben te gebruiken. Jouw vertrouwen is belangrijk voor ons en daarom delen we gegevens over jou niet met anderen tenzij we:

- *je naam en andere informatie die jou persoonlijk zou kunnen identificeren eruit hebben verwijderd."*

Beoordeel de volgende uitspraak naar aanleiding van de tekst:
Facebook mag jouw individuele profielfoto met andere partijen delen ten gunste van commerciële doeleinden.

Dit onderdeel van de vragenlijst gaat over het bewustzijn betreffende het gebruik / misbruik van jouw persoonlijke informatie. Er zijn een aantal situaties gegeven. Een aantal zijn op waarheid berust en een aantal zijn er verzonnen. Zijn de volgende uitspraken Waar of Niet waar? Als je het antwoord niet denkt te weten, gok dan niet. Vul dan Weet ik niet in. Dit onderdeel bestaat uit 5 vragen.

16. Een bedrijf dat op Facebook adverteert kan over het algemeen jou als individu een persoonlijke aanbieding doen omdat dit bedrijf jouw persoonlijke informatie (bijv. geslacht, leeftijd en hobbies) kan downloaden.

17. De volgende vraag betreft de installatie van applicaties (apps) en games over het algemeen en niet specifiek op Facebook.

Als applicatie- of game-ontwikkelaars bij de installatie van applicaties toestemming krijgen om inzicht te krijgen in jouw contactenlijst, dan betekent dit dat jij bijvoorbeeld het telefoonnummer en/of email-adres van jouw contacten weggeeft.

18. Facebook biedt samen met advertentiebedrijven de gebruikers van Facebook persoonlijke advertenties aan. Facebook gebruikt hier uitsluitend persoonlijke informatie voor dat van de profielpagina van de gebruikers afkomstig is.

19. Facebook verdient op verschillende manieren geld, bijvoorbeeld door advertenties aan te bieden. Beoordeel de volgende uitspraak:

Over het algemeen verdient Facebook ook geld door bedrijven te laten betalen wanneer zij een 'Bedrijfs-pagina' aanmaken of registeren.

20. Alle persoonlijke informatie die jij van Facebook verwijdert, wordt binnen een termijn van 30 dagen volledig van Facebook verwijderd.

Het volgende onderdeel van de vragenlijst gaat over de controle op privacy binnen Facebook. Er zijn een aantal situaties gegeven. Een aantal zijn op waarheid berust en een aantal zijn er verzonnen. Zijn de volgende uitsprakenWaar of Niet waar? Als je het antwoord niet denkt te weten, gok dan niet. Vul dan Weet ik niet in. Dit onderdeel bestaat uit 5 vragen.

21. Als ik een bericht op mijn eigen Facebook tijdlijn plaats, kan ik per bericht een specifiek persoon kiezen voor wie het bericht zichtbaar is.

22. Facebook biedt mij de mogelijkheid om berichten en foto's van anderen, waarin zij mij hebben getagt, eerst goed te keuren voordat deze op mijn tijdlijn verschijnen.

23. Facebook biedt mij de mogelijkheid om alle foto's, video's, likes en status updates die ik heb geplaatst inzichtelijk te maken. Op deze manier biedt Facebook mij een overzicht van mijn gedrag op Facebook.

24. Als ik een vriend op Facebook blokkeer met wie ik in een gedeelde groep zit, ziet deze persoon mij niet meer binnen de ledenlijst van de groep. Ook ziet hij/ zij de berichten die ik in de groep plaats niet verschijnen.

25. Facebook biedt mij de mogelijkheid om mijn Facebook profielpagina weer te geven als een ander persoon. Op deze manier kan ik mijn profielpagina bekijken vanuit het perspectief van anderen. Dit geeft mij inzicht in hoe andere gebruikers mijn profiel zien.

In dit onderdeel van de vragenlijst worden 5 afbeeldingen getoond. Op de afbeeldingen staan typische Facebook berichten van verzonnen personen. Over de personen in de afbeeldingen is geen achtergrond bekend.
Geef aan in welke mate jij per situatie het privacyrisico inschat.

Let op! Het gaat hier er niet om of jij dit Facebook bericht zelf zou plaatsen. Schat in hoeveel privacyrisico de persoon in de afbeelding loopt.

Betekenis / definitie:
*Een **privacyrisico** is: het potentiële verlies van controle over persoonlijke informatie.*

26. Andrea Paardekoper plaatst het volgende bericht op Facebook:

In welke mate beoordeel jij het privacyrisico dat Andrea Paardekoper loopt:

Q26

27. Bart Brinkhof plaatst het volgende bericht op Facebook:
In welke mate beoordeel jij het privacyrisico dat Bart Brinkhof loopt:

Q27

28. Elene van Aerle plaatst het volgende bericht op Facebook:

In welke mate beoordeel jij het privacyrisico dat Elene van Aerle loopt:

29. Cornelis Aalen plaatst het volgende bericht op Facebook:
In welke mate beoordeel jij het privacy risico dat Cornelis Aalen loopt:

Q29

30. Antonio da Silva plaatst het volgende bericht op Facebook:

In welke mate beoordeel jij het privacyrisico dat Antonio da Silva loopt:

Q28

Q30

The image below illustrates the scale presented with each question:

Het volgende onderdeel gaat over het vertrouwen dat jij hebt in bepaalde partijen. Geef aan in welke mate jij vertrouwen hebt betreffende de onderstaand genoemde uitspraken. Dit onderdeel bestaat uit het beoordelen van 12 uitspraken.

Ik vertrouw erop dat Facebook eerlijk is in de omgang met mij.
Ik vertrouw erop dat Facebook mijn privacy niet schaadt.
Ik vertrouw erop dat Facebook mijn persoonlijke data goed beschermt.
Ik vertrouw Facebook-gebruikers dat zij voorzichtig met mijn persoonlijke informatie omgaan.
Ik vertrouw Facebook-gebruikers dat zij mij niet zullen beschamen door wat ik op Facebook zet.
Ik vertrouw Facebook-gebruikers dat zij mijn persoonlijke informatie niet tegen mij zullen gebruiken
Ik vertrouw erop dat de huidige wetten en regels mijn persoonlijke informatie op Facebook voldoende beschermen.
Ik vertrouw erop dat de wet van voldoende kwaliteit is om mij gerust te stellen betreffende het gebruik van Facebook.
Ik vertrouw erop dat de huidige wetten en regels zijn opgesteld om mijn privacy (o.a. op Facebook) te beschermen.
Ik vertrouw de partijen en organisaties waar Facebook mijn persoolijke informatie mee deelt.
Ik vertrouw erop dat de partijen en organisaties waar Facebook mijn persoonlijke informatie mee deelt de waarheid vertellen betreffende het verzamelen en opslaan van de informatie.
Ik vertrouw erop dat de partijen en organisaties waar Facebook mijn persoonlijke informatie mee deelt mijn persoonlijke informatie goed beschermen.

The image below illustrates the scale presented with each statement:

In dit onderdeel van de enquête worden 3 afbeeldingen van Facebook berichten of persoonlijke informatie op Facebook getoond. Over de personen in de afbeeldingen is geen achtergrond bekend.
Geef aan in welke mate jij vertrouwen hebt in anderen betreffende het gebruik en/of misbruik van de persoonlijke gegevens. **Let op, het gaat hier dus niet over het risico dat de persoon loopt.**

Betekenis / definitie:

Mate van vertrouwen: de mate van zekerheid en geruststelling die jij hebt in anderen betreffende het gebruik van de persoonlijke gegevens.

32. Pim Huisman plaatst het volgende bericht op Facebook:

Q32

Geef aan in welke mate jij vertrouwen hebt in anderen betreffende het gebruik en/of misbruik van de persoonlijke gegevens:

Lage mate van vertrouwen 1-2-3-4-5-6-7 Hoge mate van vertrouwen

33. Kariem Gunnink plaatst het volgende bericht op Facebook:

Q33

Geef aan in welke mate jij vertrouwen hebt in anderen betreffende het gebruik en/of misbruik van de persoonlijke gegevens:

Lage mate van vertrouwen 1-2-3-4-5-6-7 Hoge mate van vertrouwen

34. De volgende gegevens worden op Facebook geplaatst:

Q34

Geef aan in welke mate jij vertrouwen hebt in anderen betreffende het gebruik en/of misbruik van de persoonlijke gegevens:

Lage mate van vertrouwen 1-2-3-4-5-6-7 Hoge mate van vertrouwen

Dit is het laatste onderdeel van de enquête. Er worden 5 afbeeldingen getoond. Geef per afbeelding aan welke eigenschappen je in welke mate herkent in de afbeelding. Er zijn geen goede of foute antwoorden. Over de personen in de afbeeldingen is geen achtergrond bekend.
Eerst wordt er gevraagd in welke mate u **voordeel** voor de plaatser van het bericht herkent.
De andere eigenschappen die je gaat beoordelen zijn **plezier**, **zelf-presentatie** en **het onderhouden van relaties**. De definities van deze eigenschappen zijn hieronder genoemd.

Betekenis / definities:

- **Voordeel:** *het profijt dat de persoon heeft door het plaatsen van het bericht (reputatie of psychologisch)*
- **Plezier:** *gevoel van blijheid*
- **Zelf-presentatie:** *op de kaart zetten van jezelf*
- **Onderhouden van relaties:** *in stand houden of verbeteren van de relatie met iemand (of meerdere personen)*

35. Lisanne van Bosgoed zet het volgende bericht op Facebook:

Q35

Geef aan in welke mate jij denkt dat het plaatsen van dit bericht effect heeft op de volgende eigenschappen voor Lisanne:

36. Tony Hallerman zet het volgende bericht op Facebook:

Q36

Geef aan in welke mate jij denkt dat het plaatsen van dit bericht effect heeft op de volgende eigenschappen voor Tony:

37. Caroliene Blauwmeer zet het volgende bericht op Facebook:

Geef aan in welke mate jij denkt dat het plaatsen van dit bericht effect heeft op de volgende eigenschappen voor Caroliene:

38. Bert Spring in 't Veld zet het volgende bericht op Facebook:

Geef aan in welke mate jij denkt dat het plaatsen van dit bericht effect heeft op de volgende eigenschappen voor Bert:

39. Marko van der Vliers zet het volgende bericht op Facebook:

Geef aan in welke mate jij denkt dat het plaatsen van dit bericht effect heeft op de volgende eigenschappen
Dit was het laatste onderdeel van het onderzoek. Ga naar de volgende pagina om het onderzoek af te ronden.

Q38

Q37

Q39

	1	2	3	4	5	6	7	
Lage mate van voordeel	○	○	○	○	○	◉	○	Hoge mate van voordeel
Lage mate van plezier	○	○	○	○	○	◉	○	Hoge mate van plezier
Lage mate van zelf-presentatie	○	○	○	○	○	○	◉	Hoge mate van zelf-presentatie
Lage mate van het onderhouden van relaties	○	○	○	○	○	○	◉	Hoge mate van het onderhouden van relaties

APPENDIX G

Anova table Benefits

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	31,807	1	31,807	37,005	,000[b]
	Residual	92,831	108	,860		
	Total	124,639	109			
2	Regression	35,939	2	17,970	21,677	,000[c]
	Residual	88,699	107	,829		
	Total	124,639	109			

a. Dependent Variable: Benefit

b. Predictors: (Constant), Enjoyment

c. Predictors: (Constant), Enjoyment, Relationship_Maintenance

APPENDIX H

Competence scale construction

In this research multiple competence scales are constructed to illustrate the competence levels of the users. Three different methods to compute the scales are applied: 1) A *normal competence scale*, 2) a *competence scale for predicted values* and 3) a *competence scale for dichotomous questions*.
The requirements for each competence scale are equal:
Criteria:

- *The increment for each competence level is equal, except the range of competence level 1;*
- *The increment is corrected for assumed maximum variance (AMV). Users that value an item with four can only vary by three points from the expert. A valuation of 4 is balanced (not high, not low).*
- *The maximum variance between an expert and an user is 6. Therefore competence level 1 has a larger maximum range than the other levels.*
- *The minimum variance of a competence level is 0.*
- *An item score of a user can not be in more than one competence level.*

Normal competence scale
The *normal competence scale* is applied for the estimation of one individual variable. The score of this variable is directly measured in via the survey and calculated with the deviations from the golden standard. The range per competence level is determined by dividing the *assumed maximum variance* (AMV) by the number of competence levels (3 / 5 = 0.6). All competence scores for the individual variables are constructed with this method. This results in the following competence levels:

Competence level	Minimum	Maximum
Competence level 5	0,0	0,6
Competence level 4	>0,6	1,2
Competence level 3	>1,2	1,8
Competence level 2	>2,4	3,00
Competence level 1	>3,00	6

Competence scale for predicted values
In chapter 4 prediction models for *Perceived Benefits, Perceived Trust* and *Perceived Privacy Risks* are presented. The predicted values are categorized as well to determine the competence levels of the users of the main variables. These values are constructed with multiple determinants and therefore the abovementioned calculations are not suitable.
The minimum, maximum and range of the competence scale depend on the model. The competence scale for Perceived Privacy Risks is constructed below as an example.
The model ($\hat{y} = 2{,}525 + 0{,}319\chi 1 + 0{,}242\chi 2$) is constructed with *assumed maximum variance (AMV)*. The AMV has a fixed value of 3. The model ($\hat{y} = 2{,}525 + 0{,}319 * 3 + 0{,}242 * 3$) results in a predicted value of 4,208, which is the maximum of the scale.

Maximum – minimum = Total Range (4,208 – 2,525 = 1,683). Increment per Competence level (IC) = Total Range / No. Competence levels = 1,683 / 5 = 0,3366

Corrected competence level perceived damage	Minimum	Maximum
Competence level 5	2,525 (Intercept)	2,861 (Intercept + 1IC)
Competence level 4	>2,861 (Intercept + 1RC)	3,197 (Intercept + 2IC)
Competence level 3	>3,197 (Intercept + 2RC)	3,533 (Intercept + 3IC)
Competence level 2	3,533 (Intercept + 3RC)	3,869 (Intercept + 4IC)
Competence level 1	3,869 (Intercept + 4RC)	5,891 (Intercept + 0,319 * 6 + 0,242 * 6)

The calculation of the maximum value for competence level 1 varies since users are able to deviate more than AMV from the expert panel.

Competence scale for dichotomous questions

The survey also contains skill questions about the Privacy Policy, Privacy Control and Awareness of Data Exploitation. The competence levels of the users for these variables are constructed with the number of correct answers instead of the deviations from the golden standard.

A competence scale for items measured with dichotomous questions therefore requires different criteria:

1. Five is maximum number correct answers;
2. Zero is the minimum number of correct answers;
3. Users that score zero or once are place in the same competence group.

Competence level	Number of correct answers
Competence level 1	0 or 1
Competence level 2	2
Competence level 3	3
Competence level 4	4
Competence level 5	5

Appendix I

Steps to select an expert

Professors, assistant professors and Ph.D.'s of the University of Twente are considered as the first pool of experts. A minimum of two experts per tested variable is required. To select an expert the following steps are undertaken:

- Select a research department within the University of Twente that has expertise regarding the variable;
- Select the members of the research group with the best fitting expertise, education, interests and research;
- Select members according to the following sequence: Professor, assistant professor, PhD. candidate
- Check the curriculum vitae to check if the expert has the required expertise;
- Scan the publications of the potential expert.

Appendix J

Expert 2 analysis

The sentiment analysis of Expert 2 is included since this expert extremely deviates from the other experts. The demographic characteristics and the interpretation of the researcher of it are given to explain the exclusion of this expert.

Gender: Female
Nationality: German
Age: 30 – 40
Survey attitude: Extreme (often uses 1 or 7 on the LIKERT scale)
Facebook account: Yes
Visiting behavior: Once a weak
Sharing behavior: Never
Research domain: Trust, Media Psychology, Social Media Analysis

The extremely deviation scores of Expert 2 are explained by the nationality and the critical attitude. The German nationality indicates higher chances of privacy importance. The critical attitude is reflected by the extreme scores on the survey.

Appendix K

Prototype of a tool

In this research several prediction models are constructed. The models are used to predict the competence levels (estimation) of the users. These models are used in the back-end of a prototype of a tool. The tool only works with the survey specific questions.

A web-tool (and Flash) is created with the original applied survey. The significant elements from the survey are used. In the conclusions a tool is recommended to illustrate the competence levels of the users. This prototype is first version of a tool that empowers users in the awareness.

The following elements are included in the prototype:

- Perceived damage estimation;
- Perceived Future Risk estimation;
- Perceived Enjoyment estimation;
- Perceived Relationship maintenance estimation;
- Privacy policy skill;
- Privacy control skill;
- Awareness of data exploitation knowledge.