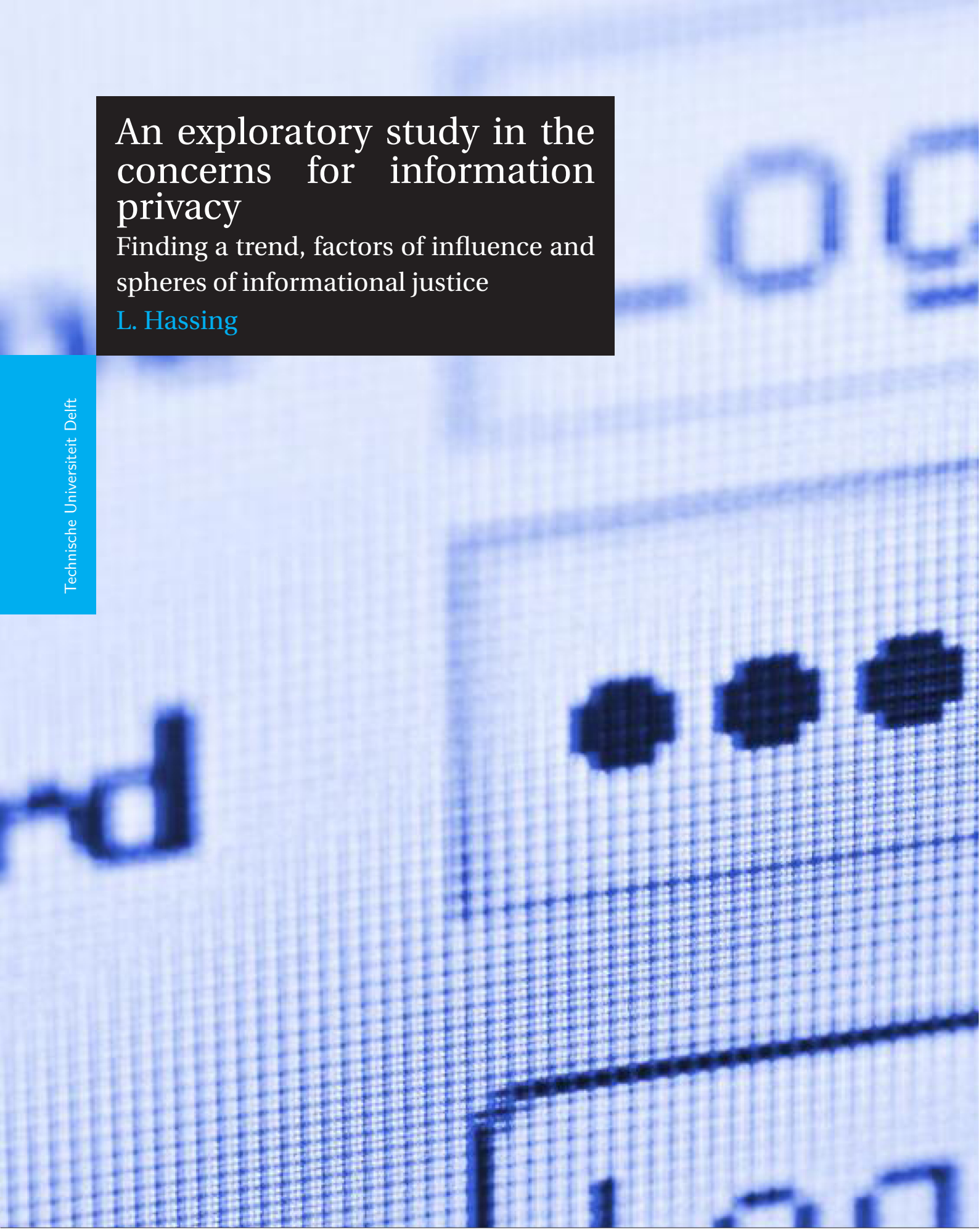
An exploratory study in the concerns for information privacy

Finding a trend, factors of influence and spheres of informational justice

L. Hassing

Technische Universiteit Delft



TU Delft Delft University of Technology

Challenge the future

An exploratory study in the concerns for information privacy

FINDING A TREND, FACTORS OF INFLUENCE AND SPHERES OF INFORMATIONAL JUSTICE

by

L. Hassing

in partial fulfilment of the requirements for the degree of

Master of Science
in Management of Technology

at the Delft University of Technology,
to be defended publicly on Tuesday January 27, 2015 at 15:00

Student number:	1508652	
Project duration:	March 19, 2014 – January 27, 2015	
Thesis committee:		
Chairman:	Prof. dr. M. J. van den Hoven	TU Delft
First supervisor:	Dr. L. Rook	TU Delft
Second supervisor:	Dr. C. I. M. Nevejan	TU Delft
Extra supervisor:	Ir. T. A. Hennis	DelftX

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft Delft University of Technology

ABSTRACT

In this research a quantitative meta-analysis was conducted on the Concern For Information Privacy construct (CFIP) and its dimensions to find a trend. This study concluded that on a global scale, people's concerns for unauthorized secondary use and improper access of their personal information have increased, and that for a general/student population in a general/online context the overall concerns for information privacy have increased in the US.

A follow-up study into the antecedents of people's concerns for information privacy concluded that MacKenzie's certainty trough applies also to people's knowledge of personal data collection and usage, and concern for information privacy. A significant conceptual model was developed to further explain this relation with respect to other determinants. Several relations between demographics and CFIP were found, for example between the highest level of education, financial stability and household income per household member and the CFIP construct. Several human values also seemed to be of significance with respect to the CFIP.

A final study tested the extent to which these findings can be explained by the spheres of informational justice theory. Evidence was found that the theory only applies to the governmental, medical and educational spheres and not to the commercial and financial spheres, indicating a clear discrepancy between the public and private domains. The salience of sub-spheres and the influence of the consequences of an exchange of personal data on its perceived appropriateness were also proven. This study led to the conclusion that public and especially private organisations should take steps to address the increasing concerns for information privacy to mitigate any adverse effects.

Frontpage image retrieved from http://static2.businessinsider.com/image/51d741c2ecad04e53a00000b-1200/now-that-you-know-how-to-harness-the-power-of-the-internet-make-sure-youre-protected-.jpg

PREFACE

Dear reader,

I would like to use this opportunity to tell the tale of how I got caught up into writing a complete thesis on the subject of information privacy concerns. Previously less captivated by the issue of information privacy, or any ethical issue for that matter, I started searching for a graduation topic on big data and quite swiftly came to the issue of privacy infringement. At first, being completely convinced by the massive potential utility of big data applications, I looked at privacy as an annoyance and a barrier for much greater technical developments. Perhaps partly due to my background in applied physics I tended to favour technological development. It felt natural to me that privacy should make way for these kinds of technical developments because that is all what I have been seeing my entire life. More internet equals less privacy, so to speak. But the more I started reading on the topic, the more it came to me that human rights should not be given up for the sake of technological innovation, especially not the fundamental and valuable right of privacy. This new knowledge has changed me and my perceptions in some way and if you continue to read this document, the same could happen to you, so read with caution!

I would also like to use this opportunity to thank everyone who has been involved in this thesis in one way or another. First I would like to thank Thieme, my contact from DelftX, for his many contributions and continuous help throughout the process. He often helped me with questionnaire items and such and even facilitated a small collaboration with researchers from Stanford. I would of course also like to thank my graduation committee: Jeroen van den Hoven, Caroline Nevejan and special thanks to Laurens Rook. Laurens was of great help when I needed him from tips on statistical analyses to detailed textual corrections.

I would also like to thank my friends and family for all of the support during the writing of this thesis. Special thanks go out to Anneriek for helping me through those lonely nights and to Adje for helping me smile even when the orange inside of the faculty had nothing to smile about.

L. Hassing
The Hague, December 2014

CONTENTS

List of Figures

List of Tables

NOMENCLATURE

Symbol	Definition
CFIP	Concern For Information Privacy
IUIPC	Internet Users' Information Privacy Concerns
GIPC	General Information Privacy Concern
IPC	Internet Privacy Concerns
μ	Mean
μ_7	Mean on a seven point Likert scale
μ_5	Mean on a five point Likert scale
$\mu_{neutral}$	Neutral middle point of the scale. (e.g. for a seven point Likert scale $\mu_{neutral} = 4$)
σ	Standard deviation
m	Slope of a linear fit
p	Significance or chance
N	Sample size
t	Difference between means divided by standard error, an indication for the relative distance between means
df	Degrees of freedom, N minus amount of parameters used in estimation
$x_{cut-off}$	Point of zero slope in parabola analysis
ρ	Spearman's correlation coefficient
R	Linear correlation coefficient or Pearson's correlation coefficient
R^2	Coefficient of determination

List of symbols and abbreviations used in this thesis.

1

INTRODUCTION

Today's world has changed tremendously since the rise of the internet. This has brought mostly good things, but unfortunately the negative consequences of these large changes are often overlooked. Critics say that the concept of privacy will have completely disappeared in the coming decades. (Rauhofer, 2008) Others, though, state that this is the moment society must choose to retain its societal values and not let them diminish for the sake of technological development. (Richards and King, 2014) Either way, the coming decade will be critical and determine how society will handle privacy in the future. For that reason, the aim of the present research is to contribute to the information privacy research field at this critical time by conducting three quantitative researches on the concerns for information privacy.

The research presented in this thesis will help to improve the currently available literature greatly on several different facets. First, and unlike any existing research, this research will investigate the changing of privacy concerns over time. It will also seek replication of some research on concerns for information privacy and add to this existing research field with new findings. Apart from this, the research presented here will be the first to quantitatively test the spheres of informational justice theory.

1.1. CONTEXT

The research field of concerns for information privacy was initiated by Smith *et al.* in 1996 with the development of the first information privacy concerns measurement tool. This scale measured the concern for information privacy (CFIP) on the basis of four dimensions: collection, unauthorized secondary use, errors and improper access. Since the development of this construct, several other constructs have been developed, such as a similar second order construct with the same dimensions (Stewart and Segars, 2002), a second order construct with the three dimensions of control, awareness and collection (Malhotra *et al.*, 2004), a two dimension construct of abuse and finding (Dinev and Hart, 2004) and a more specific unidimensional scale (Buchanan *et al.*, 2007). Despite all these developments, the original scale of Smith *et al.* with and without the adjustments of Stewart and Segars (2002) has been used most often.

Since the development of the CFIP scale of Smith *et al.*, much research has been carried out on factors or antecedents of CFIP. On an individual level, demographic factors, personality traits and personal knowledge and experience have been common targets of research. Examples of this are the positive correlation between age and CFIP (Janda and Fair, 2004) (Joinson *et al.*, 2010) (Laric *et al.*, 2009) and the relation between the Big Five personality traits and CFIP. (Korzaan and Boswell, 2008) (Junglas *et al.* 2008)

A literature research of Li (2011) presents a thorough review of this research field. Despite the quantity of

research in this field, many questions are still unanswered. For example, the influence of time has not been researched before and there has not been found any significant relation between the CFIP and education or income.

Another similar research field is also growing rapidly: research on the spheres of informational justice theory (Van den Hoven, 1999), which is quite similar to the theory of contextual integrity (Nissenbaum, 2004). Spheres refer to the different public domains of society, for example: the medical sphere, the political sphere, the commercial sphere and the educational sphere. Van den Hoven stated that:

> *"The idea of separate spheres with (information) goods and rules of access or schemes of allocation internal to them is intuitively plausible ..." (Van den Hoven, 1999, p. 148)*

This research will put this theory to the test, and find out if it can account for some of the effects observed in relation to concerns for information privacy.

1.2. MAIN RESEARCH QUESTIONS

Because a total of three studies will be conducted in this thesis, three main research questions have been developed. The main research questions are:

1. *How has the concern for information privacy developed over time?*
2. *What are important factors influencing an individual's concern for information privacy?*
3. *Is an exchange of personal data inside a sphere considered more appropriate than an exchange between different spheres?*

The following section will elaborate on how these three questions will be answered in the three corresponding studies.

1.3. SCOPE

Study 1 will consist of a longitudinal meta analysis of 20 years of research into the construct "concern for information privacy" to explore whether the opinion of society on information privacy is shifting towards a more tolerant or a more strict attitude. Many things changed over the past 20 years, so it will be very interesting to see how people's concerns for information privacy responded to this.

Study 2 will explore the current attitude of society on information privacy, and the influence of several factors on the concerns for information privacy. This research will be conducted quantitatively, using a questionnaire to gauge society's current opinion.

A lot of research has been conducted on factors which influence concerns for information privacy. These factors range from psychological determinants to behavioural intentions. Therefore a selection of factors has been made which have or have not been researched before, such as Schwartz' human values, the degree of knowledge about personal data collection and usage activities, the lack of realistic alternatives and the influence of demographics. Naturally, many factors have also been excluded from this study. The CFIP construct is a multi-level concept and the focus will be on individual level factors. So social-relational level factors, macro-environmental level factors and organisational and task environmental level factors are out of the focus of this study. Examples of individual factors which can influence someone's concern for information privacy which have been excluded from this study are power, computer anxiety and computer self-efficacy.

Study 3 will investigate whether the spheres of justice theory applies to exchanges of personal information. This will also be done quantitatively by asking people their opinion on several exchanges of personal information between organisations between different spheres or within the same sphere. This study will also

explore the existence of sub-spheres and the effect of the consequence of a personal data exchange on its perceived appropriateness. A more thorough explanation of this theory can be found in the background chapter. There will also be touched upon the influence of the data receiver, the data type and the intention of the data sender.

1.4. DOCUMENT STRUCTURE

In the following chapter the theoretical background of the present research will be explained. Here, the current state of affairs regarding information privacy concerns research and other relevant research fields will be outlined. From this theory the developed hypotheses of this research will be deduced and the conceptual models for all three studies will be presented.

Next, three different chapters will cover the three different studies. Each chapter will first discuss the methodology of the study, the research objective, sample specifications, data acquisition, questionnaire specifications and filtering methods. All the findings will be discussed in the results section and be visualized in the previously developed conceptual models.

The discussion chapter will cover the scientific and practical relevance of the present research. Here, you can also find the limitations of this research, ideas for further research and a summary of the achieved results. Finally, at the end of this report, the appendices and the bibliography are listed.

2

THEORY

In this chapter the relevant theory will be discussed and the hypotheses for the three intended studies will be presented. The hypotheses will be presented after describing each relevant theory and previous results in literature.

2.1. THE CONCEPT OF PRIVACY

Privacy is one of the fundamental human rights in society. However, technological developments in the digital era are making it increasingly difficult for society to hold on to this fundamental human right, possibly causing people's concerns for information privacy to increase.

First of all, the definition of information privacy will be discussed. A commonly accepted definition of privacy is: "the claim of individuals, groups or institutions to determine for themselves when, how, and to what extent information about them is communicated to others." (Westin, 1970, p. 7) Another definition of information privacy simply refers to the privacy of personal information and usually relates to personal data stored on computer systems. Personal information is in turn defined by the EU Data Protection Directive as any information relating directly or indirectly to an identified or identifiable natural person. Depending on the type of information, information privacy often overlaps with internet privacy, financial privacy and medical privacy.

What I would also like to highlight is the difference between informational privacy and decisional or constitutional privacy. The former relates to the two definitions presented above, whereas the latter does not describe the privacy of a person's information, but the privacy of a person's decisions. So a violation of decisional privacy occurs if one interferes with a person's (intimate) decisions. That being said, I will not consider decisional privacy in this thesis any further, but solely focus on informational privacy.

I am aware of the tremendous amount of literature on the concept of privacy, but this will not be discussed any further. The definitions presented above will be maintained in the remainder of this thesis. Apart from this, this thesis will focus especially on the *concerns for* information privacy.

2.1.1. INFORMATION PRIVACY CONCERNS AS A MULTI-LEVEL CONCEPT

Belanger and Crossler (2011) found that information privacy concerns can be conceptualized as a multi-level concept as follows:

- Individual information privacy concern: the actual concerns for information privacy of individuals.

- Group information privacy concern: "the collective concern that group members have regarding the privacy of the information the group possesses and has access to." (Bélanger and Crossler, 2011, p.1032)
- Organisational information privacy concern: "organizational information privacy concerns reflect the overall concern that organizational leaders have regarding the privacy of the information the organization possesses and has access to. Such concerns typically arise from management practices and policies." (Bélanger and Crossler, 2011, p.1033)
- Societal information privacy concern: "societal information privacy concern refers to the overall concerns citizens in societies taken as a whole have for the privacy of the information about them." (Bélanger and Crossler, 2011, p.1034)

The multi-level concept of information privacy concerns in Figure 2.1 has several remarkable elements. It first highlights the importance of individual differences and their impact on individual information privacy concern. Besides this, individual information privacy concern is influenced by the information privacy concerns of all three other levels: group, organisational and societal.

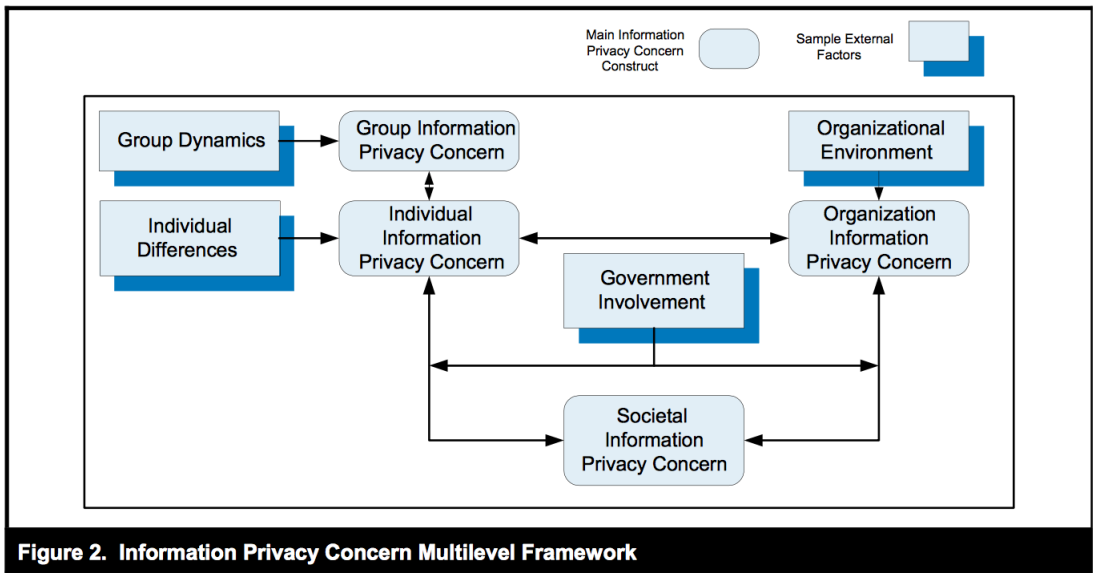


Figure 2.1: Here you can see the multi-level conceptualization as created by Belanger and Crossler in 2011.

Admittedly, privacy concerns on these more general levels are very important, but for the present study, I will focus on the lowest level. What I am interested in is people's individual privacy concerns and therefore also their individual information privacy.

In the remainder of this thesis, when using the term information privacy I will be referring to individual information privacy.

2.2. CONCERNS FOR INFORMATION PRIVACY

Current literature distinguishes several different constructs which relate to information privacy concerns. The most important and widely used are "Global Information Privacy Concerns", "Internet User's Information Privacy Concerns" and "Concern For Information Privacy", commonly referred to as GIPC, IUIPC and CFIP. A comparison of these three concepts with the definitions of the different purposes and focusses can be found in Table 2.1. Even though these concepts differ from each other, a recent study integrated the above described constructs on individual information privacy concerns into a single comprehensive concept. (Hong

and Thong, 2013) The schematic visualization of the construct and accompanying results of the model fit can be found in Figure 2.2.

	GIPC	CFIP	IUIPC
Purpose	To reflect the level of information privacy concerns in general	To reflect individuals' concerns about organizational information privacy practices	To reflect Internet users' concerns about information privacy
Focus	No particular focus	Organizations' responsibilities for the proper handling of customer information	Individuals' perceptions of fairness/justice in the context of information privacy
Context	Context-independent	Mostly offline or traditional direct marketing	Mostly online environment
Communication	Both one-way and two-way communication	Mostly one-way communication	Mostly two-way communication
Dimensions	One-dimensional construct	Collection, improper access, unauthorized secondary use and error	Collection, control and awareness of privacy practices
Representation	A single latent factor	Correlated first-order factors; Stewart and Segars (2002) argued that CFIP is better represented as a second-order factor	Second-order factor

Table 2.1: Comparison between GIPC, CFIP and IUIPC (Malhotra, Kim, and Agarwal, 2004), p.340 table 1.

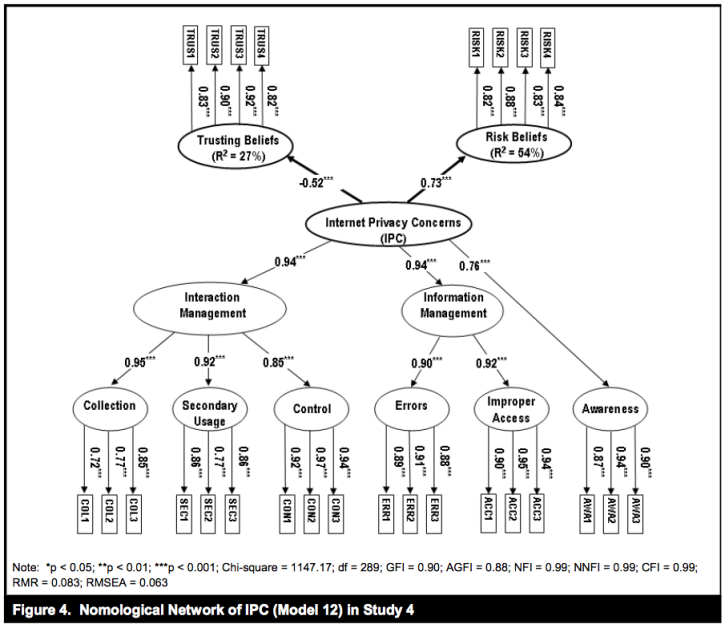


Figure 2.2: Here you can see the visualisation of the third order information privacy concern concept as developed by Hong and Thong in 2013. This model integrates all previously used concepts and claim to outperform all of them.

The IPC model describes a total of six first order factors which all originate from previous research. Five of these are affected via the second order factors "Interaction Management" and "Information Management". Interestingly, this approach renders the Internet Privacy Concerns construct a third order factor.

Although the development of this construct is very important, in the present research the focus is on the more general CFIP construct, which is not limited to internet privacy alone but instead taps into a variety of general organisational concerns. The IPC construct was also found in a late stadium of this master thesis. Because of the above reasons, the IPC construct is left out of consideration in all three studies.

2.3. THE CONCERN FOR INFORMATION PRIVACY CONSTRUCT

In this research, the choice was made to work with the construct Concern For Information Privacy, as developed in 1996 by Smith, Milberg, and Burke. This allowed for the execution of a longitudinal meta analysis of previous studies that relied on this construct to find out whether the CFIP and the four dimensions, Collection, Unauthorized secondary use, Improper access and Errors, changed over time. The CFIP construct was developed by Smith *et al.* (1996). During the development, it was soon clear that a complex and versatile mechanism underlies one's concerns for information privacy and that only a multi-dimensional construct could correctly address these concerns. A more thorough analysis of this construct can be found in the methodology section of chapter 3.

HYPOTHESES DEVELOPMENT

Due to the rise of the internet, people have grown very accustomed to sharing their personal data. This familiarization with personal data sharing which has occurred could have made people careless about their data sharing habits. But it could have also worked the other way around, making people more aware of the threats to their information privacy and increasing the concerns for information privacy. This observation leads to the following (directionless or "two-tailed") research question:

- *How has the concern for information privacy developed over time?*

This question refers to the general construct of CFIP, whereas the four corresponding sub questions refer to the four dimensions of the construct:

- How has the concern for collection of personal data developed over time?
- How has the concern for unauthorized secondary use of personal data developed over time?
- How has the concern for improper access of personal data developed over time?
- How has the concern for errors in personal data developed over time?

To test the main research question and the related four sub-questions, one hypothesis was developed for all questions:

- *H1: Individual's overall concerns for information privacy have changed over the past 20 years.*

2.4. CONCEPTUAL MODEL STUDY 1

Figure 2.3 shows the conceptual model of study 1, which describes the influence of time on the CFIP construct and the four accompanying dimensions. Of course, it is not the simple change in time that has made people less or more concerned for their information privacy, but other factors which have been changing over time are responsible for this. These factors could include the changes in culture, the rise of the internet and other technological developments. This is also visualized in the conceptual model shown here.

2.5. ANTECEDENTS OF PRIVACY CONCERNS

A lot of research focuses on the antecedents of privacy concerns. The conceptual model in Figure 2.4 illustrates how different kinds of factors can be related to the concerns for information privacy (Li, 2011).

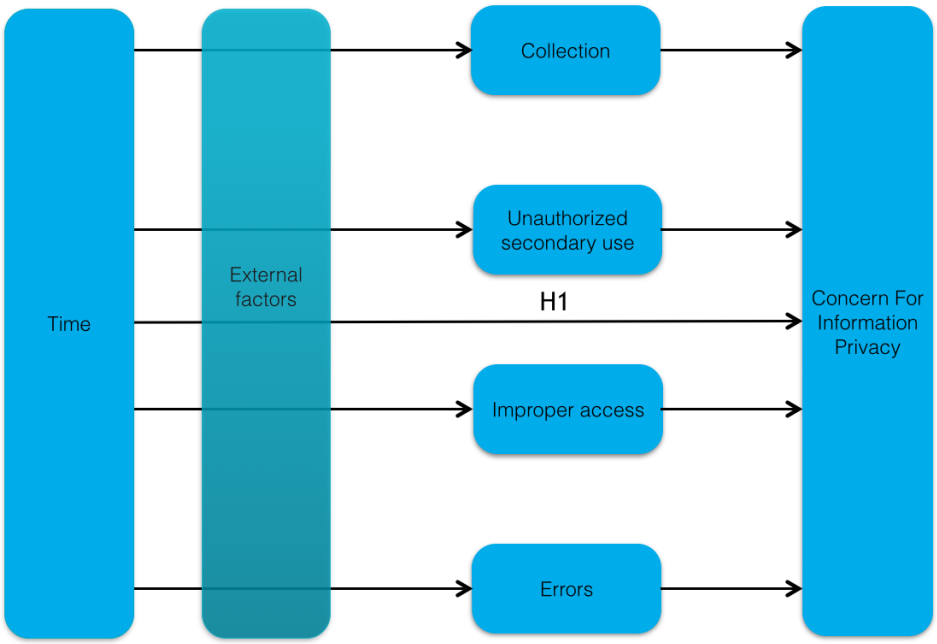


Figure 2.3: The conceptual model developed for study 1.

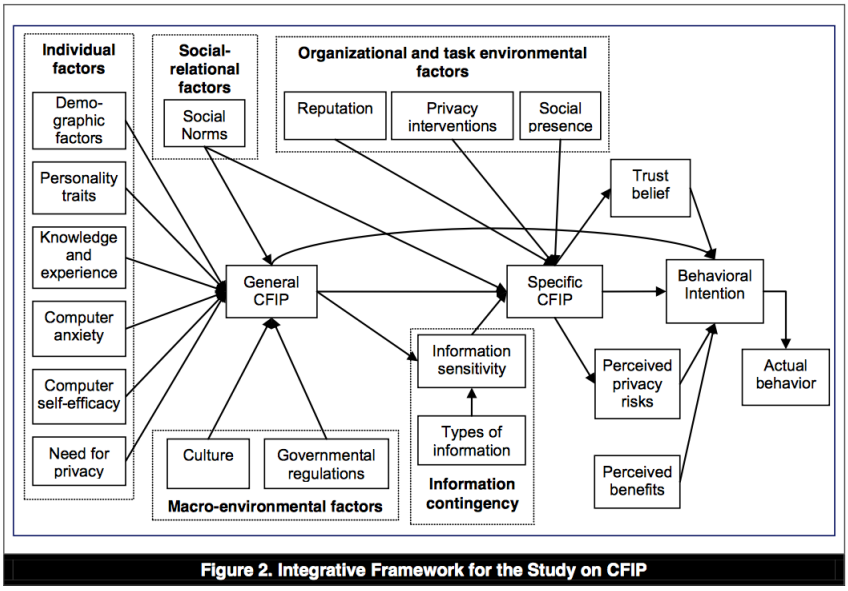


Figure 2.4: Here you can see the conceptual model of antecedents of concern for information privacy created by Yuan Li (2011) based on his literature review.

The literature review presented in this chapter will mostly focus on the individual factors which can be found on the left side of the model in Figure 2.4. The following individual factors will be further explored in the following sections:

- Demographic factors (Age, education, wealth)
- Personality traits (Human values)
- Knowledge and experience (Knowledge of collection and usage of personal data)

Apart from this the macro-environmental factors will be touched upon by discussing the influence of geographical location and the information contingency will also be touched upon by discussing the influence of different types of information on the perceived appropriateness of an exchange of personal data.

2.6. SCHWARTZ' HUMAN VALUES

Our human values define us as a species and as a society. Without these values we could not be unique, form our personalities and create our opinions. These values are the core of our existence and everything we do and feel flows from it.

Many theoretic models have been developed to describe the complete set of existing human values and how they are structured relative to one another. The most commonly used complete set of human values in the social sciences was developed by Schwartz in 1994. He proposes ten human values which can be found in Figure 2.6. In the figure resembling values are close to one another and contradicting values are depicted opposite from one another. Their similarity is based on a smallest space analysis in a two dimensional projection. This initial analysis can be found in Figure 2.5.

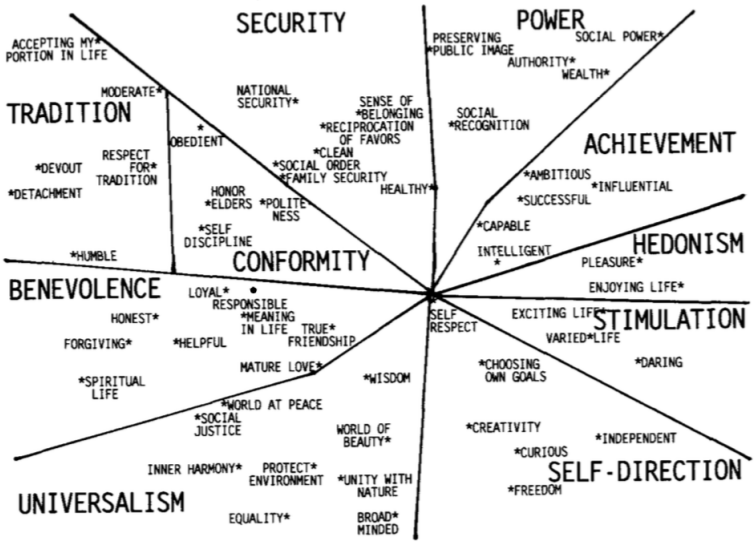


Figure 2.5: Here you can see the smallest space analysis which is the basis of Schwartz' universal human values theory (Schwartz, 1994). The scheme is developed using a questionnaire listing 56 values which was completed by 97 samples.

The theory developed by Schwartz is much more extensive and includes more complexity and diversity, but this is not further addressed in this thesis. This is done for simplicity, but could affect the outcomes of the research presented in the coming chapters.

The previous section described many kind of different factors which are antecedents of the concerns for information privacy. Quite remarkably the most true and basic antecedents, human values, have not been researched so far. Even though ones values are for a large part responsible for ones concern for information privacy.

Due to a lack of questionnaire space, only five of the ten universal human values defined by Schwartz (1994)) were explored in this thesis. These five human values are universalism, self-direction, stimulation, hedonism and security. These have been selected because they are expected to have the largest correlations with the CFIP.

Also due to the lack of questionnaire space, the original survey items of the five universal values defined by

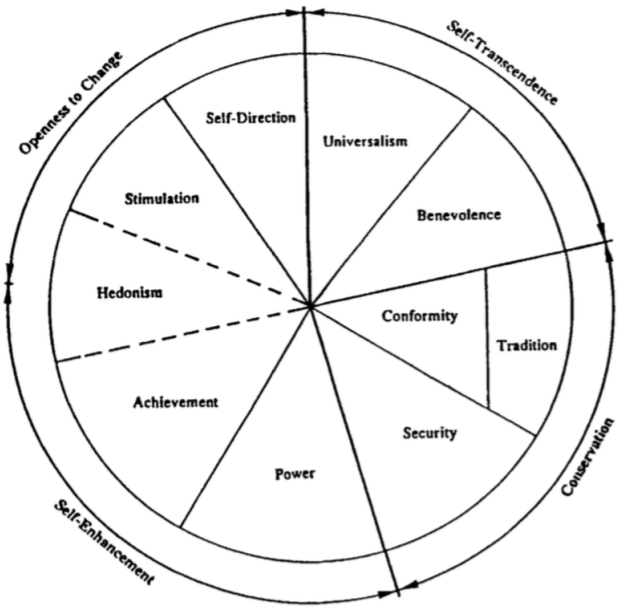


Figure 2.6: Here you can see the visualisation of Schwartz' universal human values theory (Schwartz, 1994).

Schwartz (1994)) had to be altered. Because of this alteration, the research on the relation between the CFIP and Schwartz' universal human values should be considered as an initial exploration without high reliability. Therefore, no hypotheses were developed for this relation.

2.7. DEMOGRAPHIC FACTORS AFFECTING CFIP

2.7.1. EDUCATION, AGE AND GENDER

In a recent study Li said, among other things, the following over the influence of education, age and gender on the concern for information privacy (Li, 2011):

> *"A frequently studied factor, gender, seems to exert a relatively consistent effect on privacy beliefs: except for a few studies in which insignificant effect was observed [e.g., Ji and Lieber, 2010; Yao et al., 2007], others show that women are in general more concerned about their information privacy than men [Fogel and Nehmad, 2009; Hoy and Milne, 2010; Janda and Fair, 2004; Joinson et al., 2010; Laric et al., 2009; Sheehan, 1999; Youn, 2009]. Age has a positive impact on privacy concerns in some of the studies [Janda and Fair, 2004; Joinson et al., 2010; Laric et al., 2009], but in others it influences only those without online shopping experiences [Chen et al., 2001b]; for individuals in different cultural, economic or technological environments, age may have an opposite impact on privacy concerns [Zhang et al., 2002]. Other factors, such as income and education, were not found to have a significant impact on privacy concerns across studies [Chen et al., 2001b; Ji and Lieber, 2010; Zhang et al., 2002]." (section quoted from Li, 2011, p. 460-461)*

In Tables 2.2 and 2.3 you can find some more results from previous literature on the relation between age, education and income versus the concerns for information privacy (Campbell, 1997) (Milne *et al.*, 1996). These results emphasize the significant influence of age and insignificant influence of income on the concerns about information privacy.

As for education, matters are complicated. Whereas some find no effects (see Table 2.2), others report a goodness-of-fit (see Table 2.3). Perhaps Milne *et al.* found this relation due to her Argentinian sample.

Correlation matrix Campbell 1997						
	Age		Education		Income	
	R	p	R	p	R	p
Collection	0.26	0.01	- 0.10	0.34	0.06	0.59
Access/Secondary Use	0.16	0.12	- 0.16	0.10	- 0.07	0.49
Errors	0.10	0.33	- 0.09	0.35	- 0.10	0.30

Table 2.2: These are the relevant demographic results from the research of Campbell in 1997. She found that the dimensions improper access and unauthorized secondary use were weaved together in her sample. The data of the above table has been adopted from (Campbell. 1997, p.52, Table 4).

Descriptives matrix Milne et al. 1996						
	Age		Education		Income	
	$\chi^2(6)$	p	$\chi^2(6)$	p	$\chi^2(3)$	p
Attitude towards privacy	18.7	< 0.05	13.4	< 0.05	7.49	0.058

Table 2.3: These are the relevant demographic results from the research of Milne et al. in 1996. The data of the above table has been adopted from (Milne *et al.*. 1996, p.25).

HYPOTHESES DEVELOPMENT

To shed light on this issue, the present research will try to further explore the relation between age and CFIP, find a relation between income and CFIP and find out whether education is an antecedent of CFIP. Apart from this, it will be interesting to see whether there exists a relation between someone's position in society and the individual's concern for information privacy. If so, a higher position in society is likely to correlate with a higher concern for information privacy. This leads to the following hypotheses:

- *H5A: Individuals with a higher level of education are more concerned for their information privacy.*
- *H5C: Individuals with a higher household income per household member are more concerned for their information privacy.*
- *H5D: Individuals with a higher age are more concerned for their information privacy.*
- *H5E: Individuals with a higher household income are more concerned for their information privacy.*

2.7.2. INTERNET EXPERIENCE

Another widely studied antecedent of the concern for information privacy construct is an individual's internet experience. The research of Zviran (2008) investigated this topic and found a positive relation between web usage and concern for information privacy. Also, research has been done into:

- Internet literacy
- Internet experience
- Web usage
- Use of privacy enhancing mechanism
- Web skills
- Web experience
- Internet fluency

- Internet use diversity

A three year old literature review (Li, 2011) said the following about the topic:

> *"Internet literacy (Dinev and Hart, 2005) and Internet experience (Bellman et al., 2004) were shown to have a negative impact on privacy concerns; Web usage and use of privacy enhancing mechanisms (Zviran 2008) had a positive impact; Web skills and Web experience had no impact (Janda and Fair, 2004; Zviran 2008); and Internet use fluency and Internet use diversity both had a mixed impact on privacy concerns (Yao et al., 2007; Yao and Zhang, 2008). A possible reason for the mixed results is the variety of Internet knowledge, which may have distinct roles in privacy formation. Another reason is that the relationship between general knowledge of Internet and privacy concerns may not be linear: as the knowledge of privacy issues grows, a person may become more concerned about online privacy; with further accumulation of such knowledge, the person may learn to avoid some of the privacy risks and therefore become less concerned. More efforts are needed to examine the nature of such knowledge and its impact on privacy concerns."(section quoted from Li, 2011, p. 461)*

The general factor underlying many of the above factors is how much time an individual spends on the internet every day, i.e. web use. This relatively simple factor predicts and explains all the other factors such as literacy, experience, web skills, fluency and use diversity (because this all follows from frequent use of the internet).

HYPOTHESES DEVELOPMENT

People who use the internet more often will be more accustomed to it and have a better understanding of the collection and usage of personal data by companies. From this understanding is expected to arise faith which will decrease individuals' concerns for information privacy. In this research I would like to verify the result of Zviran (2008) using the hypotheses below. Apart from this, it would also be interesting to see whether different kinds of specific internet use have an impact on people's concerns for information privacy.

- *H11A: Individuals who spend more time online will have less concerns for information privacy.*
- *H11B: Individuals who spend more time using social media will have less concerns for information privacy.*
- *H11C: Individuals who spend more time using online forums will have less concerns for information privacy.*

Although the reasoning in the previous paragraph sounds promising, a counter argument is in place. People who use the internet more often will likely have more understanding of the collection and usage of personal data by companies and they could be frightened by the current infringement of their privacy, causing them to have more concerns for information privacy. This leads to the following hypothesis:

- *H11D: Individuals who use services which require the collection of personal data more often, will have more concerns for information privacy.*

Although hypothesis H11D is similar to hypotheses H11A to H11C, there is a critical difference. Here it concerns how often people use services which require the collection of personal data, and in many cases people are forced to keep using these services because of the lack of realistic alternatives. This could cause people's concerns for information privacy to increase, so a positive correlation is expected.

2.8. MACKENZIE'S CERTAINTY TROUGH

In 1993, MacKenzie devised a relation between the involvement in knowledge production and the uncertainty in that knowledge. People close and intimately connected to the knowledge production were more unsure about their knowledge claims than those who indirectly rely on that knowledge. In the other extreme, people who were alienated from the knowledge felt that the uncertainty was much greater (MacKenzie, 1993).

> *"There seems, therefore, to be uncertainty of two quite different kinds about the facticity of missile accuracies. To stereotype, the first is the uncertainty of the alienated and those committed to an alternative weapon system: the manned bomber. The second, perhaps more surprising, form of uncertainty, is that of those closest to the heart of the production of knowledge of accuracy. Rejecting the public critics' arguments, the latter group nevertheless find in their intimacy with this process of production reasons for doubt of a more private and more limited, but nevertheless real, kind." (section quoted from MacKenzie, 1993, p. 370-371)*

A schematic visualisation of this theory can be found in Figure 2.7.

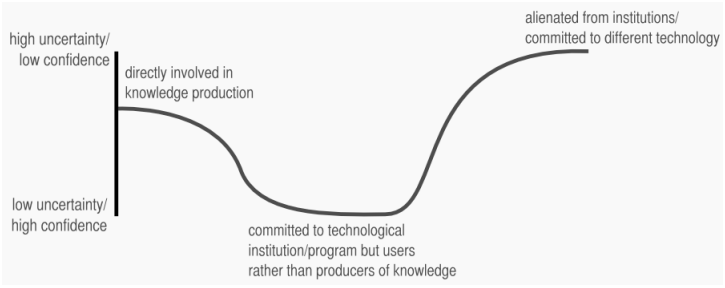


Figure 2.7: Here you can see the so called Certainty Trough, developed by MacKenzie in 1993. On the x-axis you can find the distance from knowledge production and on the y-axis you can find the degree of uncertainty.

From a more general perspective, an applied version of MacKenzie's theory can be deduced. Instead of the two variables being "distance from knowledge production" and "uncertainty/confidence", similar variables would be "knowledge of personal data collection and usage activities" and "concern for information privacy". It is interesting to find out whether the relation described by MacKenzie would also hold for these two slightly adjusted variables and to find out whether a "Privacy Concern Trough" exists.

The literature research of Li (2011) concluded the following on this topic:

> *"Personal knowledge and experience are important sources of information about privacy issues. These include general knowledge about Internet use and specific knowledge about privacy invasions. Empirical evidences of the impact of specific knowledge and experience on privacy concerns are relatively consistent, as previous experience with information misuse and disclosure [Smith et al., 1996; Okazaki et al., 2009], knowledge of media coverage on information misuse [Smith et al., 1996], and previous experience with online privacy invasion [Bansal et al., 2010; Zviran 2008] all have a positive impact on privacy concerns." (section quoted from Li, 2011, p. 461)*

So instead of looking for a linear correlation, the research in this thesis will try and find a non-monotonic relation between the knowledge of personal data collection/use activities and the concern for information privacy. Quite surprisingly, the section quoted on page 13 by Li describes an exactly opposite non-monotonic relation between general knowledge of internet and the concerns for information privacy. This non-monotonic relation was however not found in this thesis.

HYPOTHESES DEVELOPMENT

The train of thought of the previous section leads to the following hypothesis.

- *H2: Individuals with very little or very much knowledge of personal data collection and usage will have more concerns for information privacy than individuals with an average knowledge of personal data collection and usage.*

This hypothesis formally puts into words what was previously described as the Privacy Concerns Trough. An inquiry into people's knowledge of information gathering and use did reveal that the public had widely varying perceptions of the types of information accessible to marketers, as well as the accuracy of the collected personal information (Nowak and Phelps, 1992). Nevertheless, this relation was never tested before in literature and can therefore not be compared to previous results.

2.9. REALISTIC NON-TRACKING ALTERNATIVES

Recently more and more non-tracking (privacy respecting) alternatives for web services emerge such as DuckDuckGo, an alternative for Google. Apparently, the need for these non-tracking alternatives is increasing. An analysis has been conducted to map out which online services already have privacy respecting counterparts and which are free of realistic privacy respecting alternatives, this can be found in appendix H. Often these alternatives do exist, but people sometimes choose for privacy infringing services because of the better quality of the service. This phenomenon is called the privacy calculus and is discussed in the privacy paradox section. Apart from this, people are also sometimes restricted from switching to these alternatives because of network effects. Whatsapp is a excellent example of this.

So all in all, some web services have realistic alternatives but definitely not all. This fact gives rise to the need of realistic alternatives to privacy infringing web services. As far as I am aware, this need has not been analysed before in literature.

HYPOTHESES DEVELOPMENT

The following hypotheses do not describe people's concerns for information privacy, but the creation of a need which arises from these concerns. It is expected that the more concerns for information privacy people have, the more they will feel the need for realistic non-tracking alternatives to services which require the collection of their personal information.

- *H3: Individuals with higher levels of concern for information privacy will deem realistic alternatives to services which require the collection of personal information more important.*
- *H4: The relation of hypothesis 3 will be affected by the frequency of use of services which require the collection of personal information.*

The relation described in H3 is expected to be influenced by the frequency people use these services which require the collection of personal information. When people do not or rarely use services which require the collection of personal information, it would not make sense for them to suddenly feel the need for realistic non-tracking alternatives to these services. It could be argued that people who do feel the need for non-tracking alternatives have already switched to these non-tracking alternatives, but as discussed above, this is often not possible.

2.10. PRIVACY PARADOX

The privacy paradox was described as the phenomenon where an individual expresses strong privacy concerns but behaves contradictory to these concerns. For example, despite self-reported privacy concerns, some consumers still share their personal information (Gross and Acquisti, 2005)(Pavlou, 2011).

According to the privacy paradox, there is a gap between people's concern for information privacy and their actual disclosure behaviour. A theory that aims to explain this gap is called the privacy calculus. This theory states that consumers are willing to disclose their personal information as long as they benefit from it in some way or another. The most common example to illustrate this is registering for a website to be able to make use of its services.

2.10.1. WILLINGNESS TO DISCLOSE

People's willingness to disclose is quite naturally impacted directly by their concern for information privacy. Phelps *et al.* (2000) concluded: "The findings consistently reveal a strong relationship between respondents' level of concern over the ways companies use personal information and respondents' information-related beliefs and behaviours. Respondents who were very concerned were significantly more likely than other respondents to (1) believe there should be limits on how much information companies can collect from consumers, (2) believe it is wrong for companies to provide customer mailing lists to other companies or organizations, and (3) have requested that a company or organization remove their name from a mailing or telephone calling list."

This research described in this thesis was executed in cooperation with DelftX, the responsible party for the MOOC's of edX.org from the Delft University of Technology. These "Massive Open Online Courses" encompass a growing field of research, because of the newly developed research field of learning analytics and the massive possibilities of data acquisition.

That is why it would be interesting to find out how the relation between willingness to disclose and the CFIP will look like for consumers in a MOOC setting. This could perhaps be influenced by the confidential reputation of the MOOC setting and the privacy calculus theory could also interfere, since users of the MOOC are receiving a very significant service in return, namely free (and often acknowledged) education.

HYPOTHESES DEVELOPMENT

In this research I would like to test whether the relation between the concern for information privacy and willingness to disclose also holds in a MOOC setting, and if not, find out how it does look like. The following hypothesis has been developed for this purpose.

- *H9: Individuals with more concerns for information privacy will be less willing to disclose their information in a MOOC setting.*

During the testing of this relation, it will be tested whether the respondents are willing to disclose their information to *edX researchers*. It will be interesting to see whether the relation still holds if the organisation in question is a reliable educational organisation.

2.11. WEALTH AFFECTING CFIP

There has been some research on the effect of income on someone's concern for information privacy, but unfortunately, the results from these studies are inconclusive (Li, 2011).

It makes sense, though, to take wealth into account. Apart from directly looking at a relation between household income and household income per household member versus the concern for information privacy, people's financial stress, financial stability and position in society will also be investigated.

First, having high financial stability indicates a certain level of wealth. Drawing from research by Maslow (1943) it is expected that people with more wealth will be more concerned for their information privacy. This can be understood by using Maslow's hierarchy of needs theory which is displayed in Figure 2.8 (Maslow, 1943).

2.11.1. HIERARCHY OF NEEDS

The hierarchy of needs theory was developed by Maslow in 1943 in his paper "A Theory of Human Motivation". The theory is visualized triangularly in Figure 2.8, and states a certain need can only dominate the human organism when all other needs lower in the pyramid are satisfied.

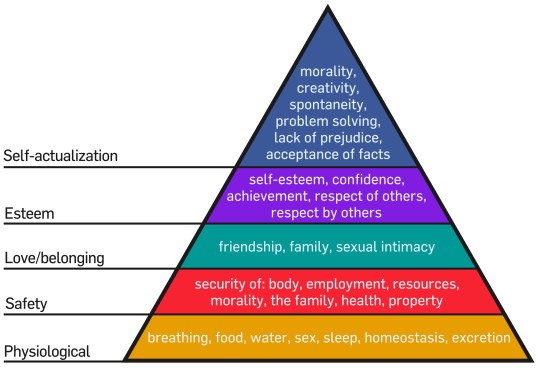


Figure 2.8: Maslow's Hierarchy of Needs Theory in a triangle shaped visual representation. (Maslow, 1943)

At the very bottom of the pyramid are the physiological needs. These are the requirements for human survival such as food, water and shelter. Directly above this are the safety needs. These needs include: personal security, financial security, health and well-being and a safety net against accidents/illness and their adverse impacts. In the next section I will discuss the notion of privacy in this category of needs. The middle category is love and belonging. This category relates to the need to maintain emotionally significant relationships such as friendship and family and the need for intimacy. The second last category is esteem, which is the need to feel respected and to have self-esteem and self-respect. At the very top of the pyramid is the need for self-actualization, which is the need of a person to realize its full potential, to accomplish everything that one can. This need can only dominate with all other needs being satisfied. The theory developed by Maslow includes more complexity and diversity, but this is not further addressed in this thesis. This is done for simplicity, so the theory of Maslow can be used to explain one aspect of the relation between wealth and the CFIP construct.

From the above theory can be seen that physiological needs like food and water are the most basic needs which have to be fulfilled first before one can worry about higher needs such as safety. Although the theory states that financial security is a need in the "safety" category, I think that having enough money or being financially stable facilitates the physiological needs. Therefore, the need for financial stability should be in the very bottom of the safety category. The safety of an individuals' personal information is therefore higher in the pyramid than the need for financial stability, meaning that the need for the safety of an individuals' personal information can only dominate if the need for financial stability is satisfied. So it is expected that financial stability is positively correlated with concerns for information privacy. This leads to the following hypotheses:

HYPOTHESES DEVELOPMENT

- *H10A: Financially independent individuals with more financial stress will have more concerns for information privacy.*
- *H10B: Financially independent individuals with more financial stability will have more concerns for information privacy.*
- *H10C: Individuals who are financially independent will have more concerns for information privacy than individuals who are financially dependent.*
- *H5B: Individuals with a better position in society are more concerned for their information privacy.*

The above hypotheses will explore the influence of financial stress, stability and independence on the concerns for information privacy. The financial independence variable will serve as a selection criteria, because financially dependent individuals will not truly experience financial stress and their financial stability is not the result of their own actions. Thus only financially independent individuals will be taken into account.

The third hypothesis H10C very much resembles hypothesis H5A, because age will be the deciding underlying factor which explains the discrepancy in information privacy concerns between financially dependent and financially independent individuals.

2.12. GEOGRAPHICAL FACTORS AFFECTING CFIP

Research by Bellman *et al.* (2004) discovered that cultural values do have an influence on consumers' concerns about information privacy, an observation which was also concluded by Milberg *et al.* (2000). The analysis of specific cultural values is left outside of the scope of this research, but there will be touched upon the impact of geographical location.

The geographical location is an interesting factor, because "more dramatic differences could be expected in a sample drawn from the general populations of the countries." (Milberg *et al.*, 1995, p. 72).

HYPOTHESIS DEVELOPMENT

To investigate the impact of geographical location on the concerns for information privacy the following hypothesis has been developed.

- *H12: Levels of concerns for information privacy will differ across countries.*

2.13. CONCEPTUAL MODELS STUDY 2

A visualization of all of the above hypotheses can be found the conceptual models in Figures 2.9 and 2.10. The model has been split, because the size of the model could otherwise create a lack of transparency. Also, the relations are only visualized with respect to the general concern for information privacy construct. Note though, that the relations with the four different dimensions, Collection, Unauthorized Secondary Use, Improper Access and Errors are also tested, but not visualized in the model since this would also cause clutter in the model.

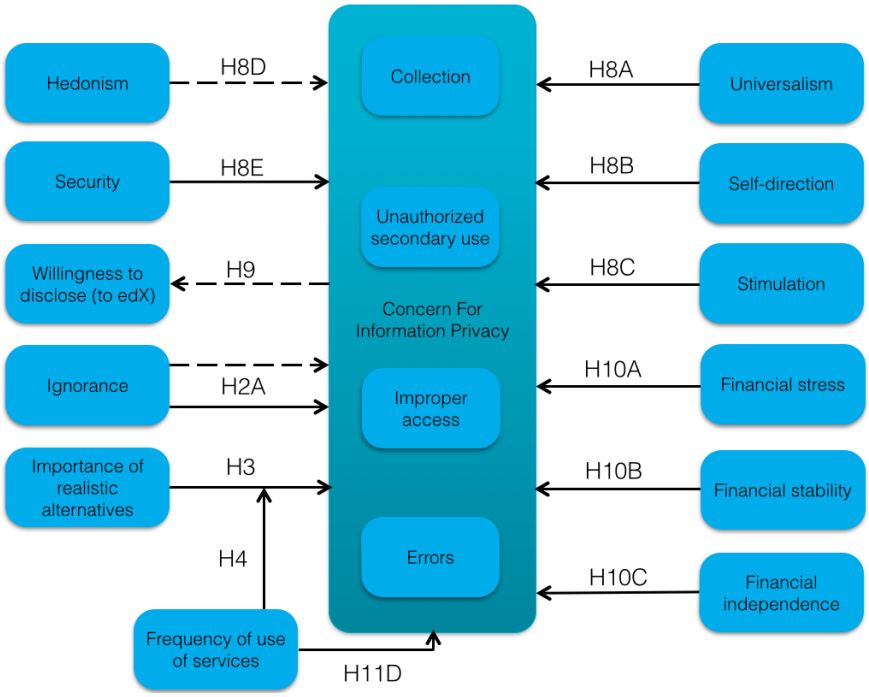


Figure 2.9: The first conceptual model developed for study 2. Dashed lines indicate a negative relation and solid lines indicate a positive relation.

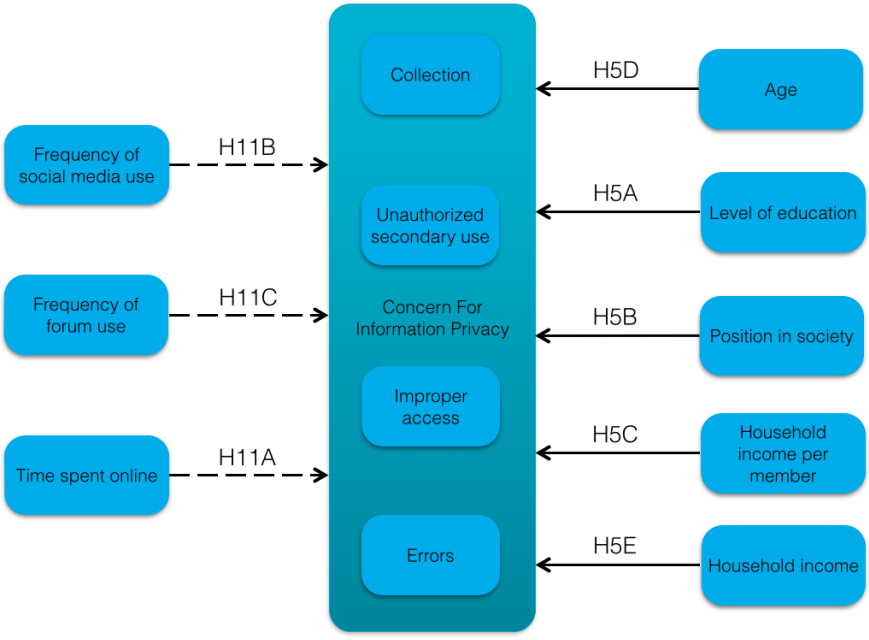


Figure 2.10: The second conceptual model developed for study 2. Dashed lines indicate a negative relation and solid lines indicate a positive relation.

2.14. SPHERES OF JUSTICE

The original spheres of justice theory was developed by Michael Walzer in 1983 in his book called "Spheres of justice, a defense of pluralism and quality." Spheres refer to the different public domains of society, for example: the medical sphere, the political sphere, the commercial sphere and the educational sphere. All these

spheres have different goods, and these goods are allocated differently, e.g. medical treatment is allocated on the basis of need, political office on the basis of democratic election and money on the basis of free exchange (Van Den Hoven, 2008). The theory states that different spheres contain different goods allocated by means of different allocation criteria or distributive practices (Van Den Hoven, 2008). According to this theory, people especially find it offensive when:

1. Internal goods in a sphere are allocated on the basis of the distributive or normative logic associated with another sphere;
2. The transfer of goods across the boundaries of separate spheres;
3. The dominance and tyranny of some goods over others.

In either of these instances, people will experience injustice. The above three forms of injustice can be prevented by applying the "art of separation" and by putting blocked exchanges into place. (See Figure 2.11).

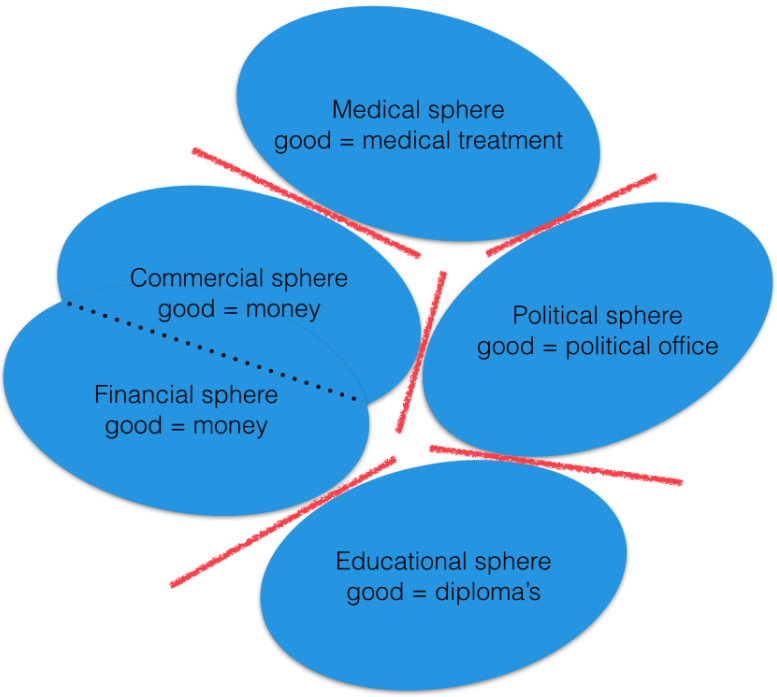


Figure 2.11: Here you can see a schematic visualisation of the Spheres of Justice theory, developed by Michael Walzer in 1983. The red lines indicate the blocked exchanges and the commercial sphere is split up into "commercial" and "financial". This is done because I wanted to make a clear distinction between banks and other companies in study 3.

2.14.1. SPHERES OF INFORMATIONAL JUSTICE

It seems theoretically plausible to assume that these forms of injustice also apply to exchanges of personal information. Although personal information is not a good which particularly associates with a specific sphere, intuitively it can be imagined that exchanges of personal data inside the same sphere are more just, or more appropriate, than exchanges of personal data between different spheres. For example: If a hospital shares your medical records with another hospital, no offence will be taken. But when a hospital shares these data with commercial companies it is considered inappropriate. This application of Walzer's Spheres of Justice on Information privacy was first developed by Van den Hoven in 1999. This theory will serve as a guideline in study 3 and the goal of study 3 is to prove and add to this theory, but in the following section a similar complementing theory will be discussed.

2.14.2. CONTEXTUAL INTEGRITY

I would like to introduce this theory by presenting a quote of the initial paper:

> *"A central tenet of contextual integrity is that there are no arenas of life not governed by norms of information flow, no information or spheres of life for which "anything goes." Almost everything—things that we do, events that occur, transactions that take place—happens in a context not only of place but of politics, convention, and cultural expectation. These contexts can be as sweepingly defined as, say, spheres of life such as education, politics, and the marketplace or as finely drawn as the conventional routines of visiting the dentist, attending a family wedding, or interviewing for a job." (Nissenbaum, 2004, p. 119)*

So how we apply the concept of privacy to the activities of our daily lives depends greatly on context. And this context can be anything, from the larger spheres of informational justice initially defined by Van den Hoven (1999), to for example the type of relation you have with your siblings. This broad definition of contextual integrity makes the concept easy to use in explaining past events, but difficult to use in predicting new situations. This is caused by the unclear, overly broad definition of a context.

What lacks in this theory is a comprehensive framework explaining why these contexts exist, how they are formed and why the allocation criteria and appropriateness differ between contexts. The previously discussed spheres of informational justice theory does present this explaining framework and this is why the spheres of informational justice theory is used in the remainder of this thesis.

Note that the theory of contextual integrity does not oppose the spheres of informational justice theory, but is one hundred percent complementary. Nagenborg (2009) has attempted to link privacy as contextual integrity, the ethical design of infrastructures and Walzer's "art of separation" (the basis of the spheres of informational justice theory). He states that *"Bringing together "contextual integrity" and the "art of separation" enables us to base reflections upon information flows in and between institutions on Walzer's work" (Nagenborg, 2009, p.178)*. However, the large amount of challenges he also presents, is an indication of the large amount of research which is still required on this subject.

2.14.3. PRESENCE OF SUB-SPHERES

Past literature has defined an "informational sphere" which contains several "sub-spheres" of its own, which are also separated by the art of separation. (Nagenborg, 2009) I think however that it is more clear to see personal data as an additional social good, which can be exchanged by different specific allocation criteria in the previously defined original spheres of Walzer. Similar to how Walzer discussed privacy as a social good. So to be clear; when I talk about informational spheres, literature has defined these as "sub-spheres". And what I call informational sub-spheres is not used in literature before, but they would have been called "sub-sub-spheres".

What I would like to add to this theory of Walzer is the notion of separate domains inside spheres, which I will call sub-spheres. These sub-spheres exist on all kinds of different levels, and inside these sub-spheres, the sharing of personal data is considered even more appropriate than in the overall sphere. An example would be a university sharing your personal performance data. If this would be within the university with other teachers, or even with collaborating nearby universities, this would be considered more appropriate than sharing the same information with a foreign strange university, although this is considered the same sphere by Walzer.

Another educational example could be from the teachers' perspective when considering the sharing of performance data inside a university. The teachers would find it appropriate for them to see the performance data of their students, since this a great feedback mechanism and comparisons between teachers could significantly improve the quality of the education. The teachers would however find it inappropriate for their

dean to see the same data, because the performance data of their students is in some way also the performance data of the respective teachers and this could cause the teachers to be frightened by the power of their dean and the possible consequences. In this example the different hierarchical layers create different informational sub-spheres.

So these sub-spheres can be created on the basis of hierarchy on an organisational level, but could also for example be created on the basis of affinity or hostility on a social level.

It should be noted that the above paragraphs are intended as an addition to the spheres of informational justice theory by Van den Hoven (1999), but there is a large overlap here with the contextual integrity theory by Nissenbaum (2004). Perhaps this line of reasoning could be a way of integrating these two complementary theories.

2.14.4. INFLUENCE OF CONSEQUENCE

In a perfectly rational world, actions should be valued by their true intentions, but we live in a world built by humans where consequences are sought to be linked back to their causing actions. Despite this, claims have been made that informational cross-contaminations would be considered unjust, independent of the consequences (Van Den Hoven, 2008). It thus follows that an informational cross-contamination, would always be deemed inappropriate, regardless of the conditions.

I however think that people inherently seek to link these consequences to the responsible causing action and judge the justness of the situation based on both action and consequence. The Dutch legal system is a good example of this. A specific action could cause involuntary manslaughter, whilst the same action might also cause nothing. Both scenarios however are not considered equally just, as only one will result in punishment.

HYPOTHESES DEVELOPMENT

The line of thought of the previous section originates from a book section from 1999 called "Privacy or informational injustice" (Van den Hoven, 1999), and leads to the following hypotheses:

- *H13: The average appropriateness of personal data exchanges between spheres is lower (less appropriate) than the average appropriateness of personal data exchanges in the same sphere.*
- *H15: The average appropriateness of personal data exchanges in sub-spheres is even higher (more appropriate) than the average appropriateness of personal data exchange in normal spheres.*
- *H14: The average appropriateness of personal data exchanges is dependent on the consequence of the exchange, both in and between spheres. From least to most appropriate: personal negative consequence, societal positive consequence, personal positive consequence.*

2.15. CONCEPTUAL MODELS STUDY 3

In Figures 2.12 to 2.14 you can find the conceptual models which were developed to test the above developed hypotheses.

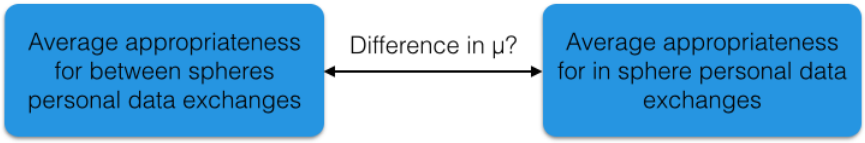


Figure 2.12: The first conceptual model developed for study 3. The goal of this model is to prove the basis of the spheres of informational justice theory, namely that intra-sphere exchanges of personal data are considered more appropriate than inter-sphere exchanges of personal data.

Figure 2.13: The second conceptual model developed for study 3. The goal of this model is to prove that the consequences of a personal data exchange influence the perceived appropriateness.

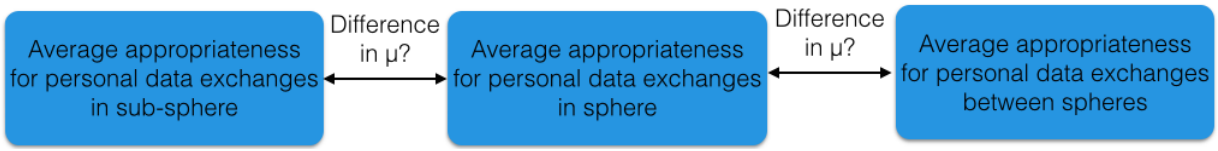


Figure 2.14: The third conceptual model developed for study 3. This model overlaps with the model presented in Figure 2.12, but the main focus of this part of the research is to prove the existence of "sub-spheres".

2.16. WORKING WITH META-CONCEPTS

This chapter has presented many different meta-concepts, and most are of high complexity. Examples are Maslow's hierarchy of needs theory, Schwartz' universal human values theory and Walzer's Spheres of Justice theory. When working with these concepts, they are often simplified and their complexity is not incorporated into the research designs. Although this could affect the outcomes of the research presented in the following chapters, this was often necessary to be able to say something meaningful about the results. This does however not mean that the complexity and diversity of these meta-concepts are not taken into account.

3

STUDY 1

3.1. METHODOLOGY

As discussed previously, current developments in technology could change the way people think about information privacy. This could also impact people's Concern For Information Privacy. Study 1 explored this thought by performing a longitudinal meta analysis of the appreciation for the Concern For Information Privacy construct over time and its four dimensions over time: Collection, Errors, Unauthorized Secondary Use and Improper Access.

3.1.1. RESEARCH OBJECTIVE

The objective of study 1 was to gain insight into the course of the Concerns For Information Privacy over the past 20 years. Much has happened the past two decades, and society is now at a crucial point in which it must decide whether to hold on to one of its fundamental values or abandon it for the sake of technological development.

"Critically, if we fail to balance the human values that we care about, like privacy, confidentiality, transparency, identity and free choice with the compelling uses of big data, our Big Data Society risks abandoning these values for the sake of innovation and expediency." (Richards and King, 2014, p. 395)

To measure how these concerns have been evolving, I chose to execute a quantitative longitudinal meta analysis. This method was selected, because it is a most objective way to truly find out how the concerns have been changing. The data of relevant scientific papers of the past 20 years were analysed to find a trend in the magnitude of the Concerns For Information Privacy. A questionnaire was found which suited this job perfectly, namely the questionnaire of Smith *et al.* from 1996. This questionnaire was published in 1996, which leaves a time span of 18 years, and it has been used thoroughly in the literature on privacy concerns. It does not solely measure the Concern For Information Privacy construct, but also four accompanying dimensions: Collection, Unauthorized Secondary Use, Improper Access and Errors. So using this questionnaire for the longitudinal meta analysis also allowed to gain insight into the development of these separate dimensions.

3.1.2. THE CONCERN FOR INFORMATION PRIVACY CONSTRUCT

In this research, the choice was made to work with the Concern For Information Privacy construct, as developed in 1996 by Smith, Milberg, and Burke, to execute a longitudinal meta analysis.

During the development of the CFIP construct, it was soon clear that a complex and versatile mechanism underlies one's Concerns For Information Privacy and that only a multi-dimensional construct could correctly model these concerns. The dimensions ultimately formed are: Collection, Unauthorized Secondary Use, Improper Access and Errors. Considering the frequent use of the CFIP construct, a thorough elaboration is in place. The separate dimensions of the concept are discussed below and a visual representation of the CFIP model can be found in Figure 3.1.

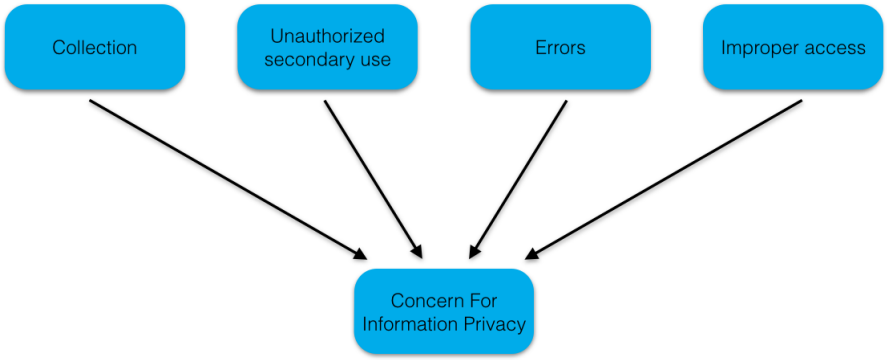


Figure 3.1: Here you can see the visualisation of CFIP model developed by Smith, Milberg and Burke in 1996.

COLLECTION

This is the most frequently occurring Concern For Information Privacy that people have. This concern is created directly from the feeling of privacy infringement, caused by the knowledge that your personal information is being collected by organisations. Purely this acquisition and ownership of an individual's personal information by an organisation is what causes the concerns for Collection, so it does not address a specific action that worries an individual.

UNAUTHORIZED SECONDARY USE

The Unauthorized Secondary Use dimension refers to both internal and external Unauthorized Secondary Use. An internal example is "sugging", in which data are initially collected for research but are later used for marketing purposes (Smith *et al.*, 1996). This often also occurs externally, when companies sell people's personal data to third parties.

People are concerned by this, because this causes them to lose control over their personal information. Even if an individual only shares their personal data with a limited number of organisations of his/her choosing, the control over the data is lost and he/she can not know where it will end up. This causes Concerns For Information Privacy.

IMPROPER ACCESS

This dimension resembles the Unauthorized Secondary Use dimension, especially internal Unauthorized Secondary Use. The slight difference with internal Unauthorized Secondary Use is that with Improper Access, the focus is not on using the data, but solely the act of accessing. Also, the Improper Access dimension focusses on the internal policy of organisations on who can access which personal information. Although this does not distribute an individual's personal data outside the company, the personal data does end up in the hands of people which should not have access to it. This fact could cause Concerns For Information Privacy.

ERRORS

The last dimension of the Concern For Information Privacy construct is the concern for Errors in personal data. Erroneous data can often cause inconvenience and problems, because of their static nature in a dynamic world. Although organisations have been looking to solve these problems the past thirty years, it seems as if handling personal data will inherently be accompanied by occasional errors.

COMBINING DATA

In early literature the combination of data was acknowledged as a potential privacy concern by Smith *et al.* (1996), but was not included into the Concern For Information Privacy construct. Anonymisation of data is a growing topic. Perhaps this concern for combing data will increase in the coming decades, since anonymised data can be traced back to its owner in combination with the correct data. Proving the possible increase of this lost dimension could be an interesting research topic, but this is left out of the scope of this thesis.

STRENGTHS AND WEAKNESSES

An advantage of using the Concern For Information Privacy construct of Smith *et al.* was that it has been formulated in an organisational context. This leaved room for the respondent to fill out the measurement instrument in his/her own interpretation. Now, technological development caused almost all sharing of personal data to be via the internet. But because respondents could interpret the organisational context automatically as an online context, the measurement item is still valid, 18 years after its creation.

The fact that the construct is fairly aged can be seen as a strength and a weakness. Its strength was that this allows for a longitudinal analysis and different studies to be compared with one another. But it could also have caused the construct to become out of date, since privacy concerns have been changing tremendously.

3.1.3. DATA ACQUISITION

The search mechanism for this longitudinal meta analysis was as follows: the goal was at first to accumulate as much papers which contained the required data as possible. After this period of divergence, the convergence/filtering phase began. Here it was important to double-check the accuracy of the acquired data and assess whether the data was suited for the analysis.

The search method mostly included using Google Scholar and Scopus to find scientific papers online according to three predefined lists which have been found. These lists were very conveniently stumbled upon during the execution of a personal literature research. They were all included in papers which originate from prestigious journals, namely: Management Information Systems Quarterly (MISQ), Communications of the Association for Information Systems and Sprouts: Working Papers on Information Systems. You can find a copy of these lists in Appendix F. Apart from these lists, several search queries were used in both Google Scholar and Scopus to find as many papers containing data from the questionnaire of Smith *et al.* as possible. A search queries which was often used was: "Concern For Information Privacy questionnaire smith 1996". But to expand the reach of this meta analysis, the query was often altered to find additional results.

3.1.4. FILTERING

When a relevant paper was found it was first verified whether it included the required data. This data was the mean and standard deviation of the Concern For Information Privacy construct and the means and standard deviations of the four dimensions of the Concern For Information Privacy construct, namely Collection, Unauthorized Secondary Use, Improper Access and Errors. Often data based on other scales was also found and saved, such us the Internet Users' Information Privacy Concerns scale of Malhotra *et al.* (2004) and the

Internet Privacy Concerns scale of Dinev and Hart (2004) (Not to be confused with the Internet Privacy Concerns scale of Hong and Thong (2013)). Nine data points were found for the Internet Privacy Concerns scale developed by Dinev and Hart, and only four data points for the Internet Users' Information Privacy Concerns scale developed by Malhotra *et al.*. Both sample sizes were much smaller than the sample size of the Concerns For Information Privacy construct, so data of the constructs of Malhotra *et al.* and Dinev and Hart were not used in any further analysis because of the superior sample size and larger time span of the CFIP construct.

Only data points which were measured by the questionnaire of Smith et al. *from 1996 were included into the data set of study 1.*

It often occurred that a certain paper did use the correct questionnaire of Smith *et al.*, but did not report the means and standard deviations of the Concern For Information Privacy construct. In this instance, the respective author was contacted via email with the question whether or not the author was willing to share this data. Unfortunately, this only paid off in two instances.

Apart from the means and standard deviations of the CFIP construct and its dimensions, some other information was extracted from the selected papers which was crucial in assessing the quality of the data point. It concerns the:

- research context;
- population size (N);
- population geographical location;
- population type (e.g. students;
- date of survey completion;
- scale specifications.

These variables helped decide which papers to include into the analysis during the filtering process. Before filtering a data set of $N = 65$ data points was collected. After filtering however, $N = 35$ clean data points were left, which all measured the Concern For Information Privacy construct correctly. You can find the table reporting the data set after filtering in appendix D in Table D.1. The data in Figure D.1 was used for the analyses in the following section.

After the data collection, it turned out that there was no statistical method to use the information provided by the standard deviations of the CFIP construct. Therefore, all collected standard deviations were used in the analysis and the collected means served as the final data points.

3.1.5. SEPARATION OF THE DATA SET

Previous research showed that research context, information type (Culnan, 1993), geographical location (Milberg, Burke, Smith, and Kallman, 1995) and population type (Bellman, Johnson, Kobrin, and Lohse, 2004) are important factors which influence people's Concern For Information Privacy. Here, geographical location and population type are direct predictors with cultural values being the indirect influencing factor.

The variable I was interested in for this study was the time the survey was completed. Ideally, I would like to keep all other variables constant to examine the pure relation between time and Concern For Information Privacy. Fortunately the amount of data points $N = 35$, just allowed for this filtering to create these ideal conditions. Most data points originated from the US, had a student or general broad population and were formulated in a general or internet context. So the three filter criteria which were used throughout this study were:

- Geographical location = United States ($N = 18$)
- Population type = field/general or student sample ($N = 21$)
- Research context = general or online ($N = 27$)

The student and field/general sample population types were assumed to be very much alike, and were therefore combined in one filter criterion. This was also the case for the general and online research contexts; these are much alike since people likely associate concerns about organisational practices automatically with online applications. And even though the US is a very large country, with differences in cultures in different geographical areas, the legislation on information privacy is the same countrywide. Therefore this functioned nicely as a selection criterion.

If all data points were to be filtered by the above three criteria, this would leave *just nine cases* for the analysis. In the results chapter it can be seen that this is enough to find significant correlations. The amount of cases left after filtering by the three selection criteria separately can be found in the above list. It was chosen to first analyse the data set with no selection criteria, than with all three selection criteria simultaneously and finally with each selection criterion separately. This allowed for a clear insight into the versatility of the data and clarified the impact of applying these different selection criteria separately.

3.1.6. SCALE TRANSFORMATIONS

For some reason, many authors chose to reduce the scale of the survey of Smith *et al.* from a seven point Likert scale to a five point Likert scale. In this case, the data point was adjusted according to the following formula:

$$\mu_7 = \frac{3}{2} \cdot \mu_5 - \frac{1}{2} \tag{3.1}$$

Where μ_5 is the original mean on the five point Likert scale and μ_7 is the adjusted mean on a seven point Likert scale. Formula 3.1 linearly projects the five point scale on the seven point scale, which was the scale that will be used throughout this thesis. This was chosen, and not the other way around, because most papers reported their data on the seven point scale. In conventional statistics this is not allowed, since the scales are ordinal and not ratio. But it was assumed that the impact of this transformation is negligible. A research from Dawes (2008) supports this assumption. In his research, different respondents were asked the same questions on different scales. The results showed that the rescaled five point Likert scale and rescaled seven point Likert scale were almost identical in terms of mean, standard deviation and skewness. Rescaling a ten point Likert scale however did change the mean of a construct, since it did not contain a neutral middle option. Therefore ten point scale data points (Tsarenko and Tojib, 2009) and four point scale data points (Culnan, 1995; Awad and Krishnan, 2006) were left out of the analysis.

3.2. RESULTS

The main research question in study 1 was: *How has the Concern For Information Privacy developed over time?*

This question referred to the general construct of CFIP, whereas the following four sub-questions refer to the four dimensions of the construct:

- How has the concern for Collection of personal data developed over time?
- How has the concern for Unauthorized Secondary Use of personal data developed over time?
- How has the concern for Improper Access of personal data developed over time?
- How has the concern for Errors in personal data developed over time?

These questions are answered in this section by testing the hypothesis H1, "Individual's overall Concerns For Information Privacy have changed over the past 20 years."

3.2.1. MAIN ANALYSIS

The full data set which was acquired consisted of $N = 35$ data points. First an analysis was done on all data points, even though these differed in geographical location, population type and research context. The analysis of the separate selections based on the previously defined selection criteria can be found at the end of this section. The data set was analysed in SPSS 22 to find a correlation between time and Concern For Information Privacy and it's dimensions. Bi-variate correlation analyses were used to test the magnitudes and the significances of the relations and scatterplots were computed to visualise the relations. These scatterplots can be found below.

Please note that Spearman's correlation coefficient (ρ) was used in all the bi-variate correlation analyses of this entire thesis. Even though this correlation coefficient is based on ranking instead of linearity, all scatterplots in this thesis were fitted with a linear approximation. These linear approximations only serve as visual support and were not used or taken into account in any of the correlation analyses!

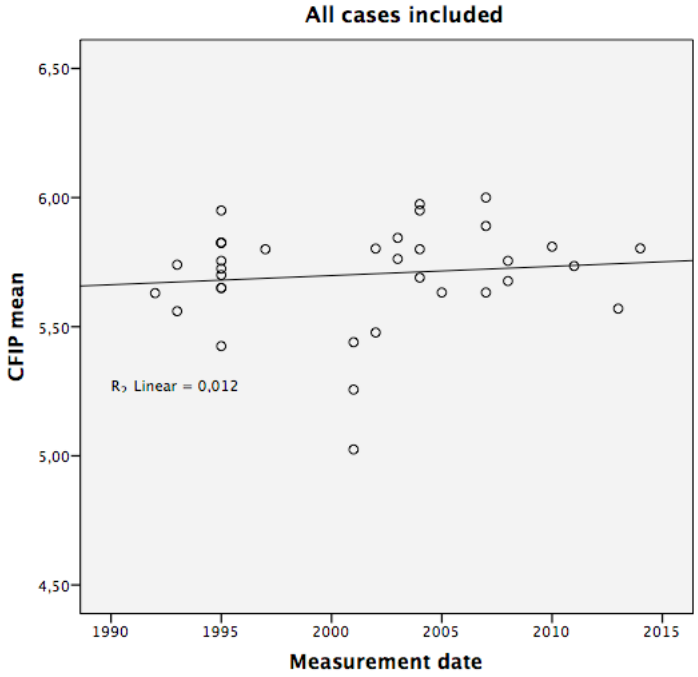


Figure 3.2: Here you can see the course of the Concern For Information Privacy over time. The concerns seemed to be somewhat constant over time.

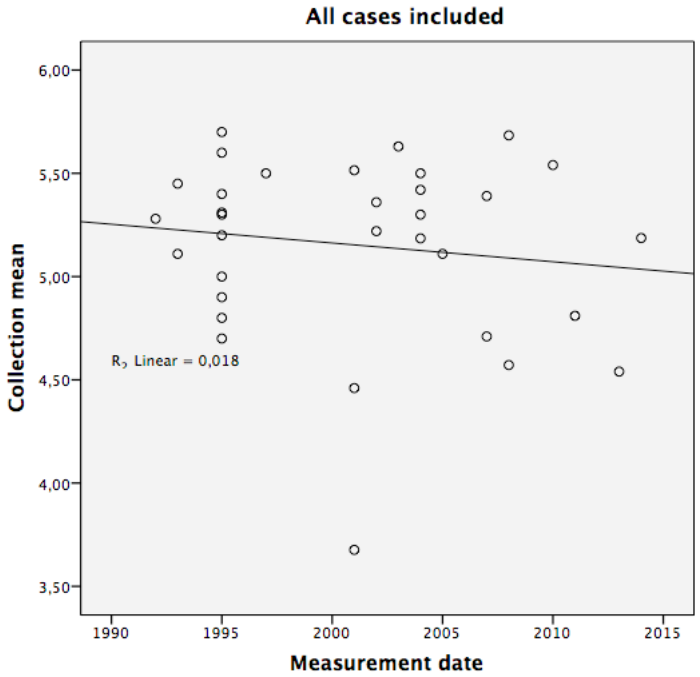


Figure 3.3: Here you can see the course of the Collection dimension over time. The concerns for this dimension seemed to be somewhat constant over time.

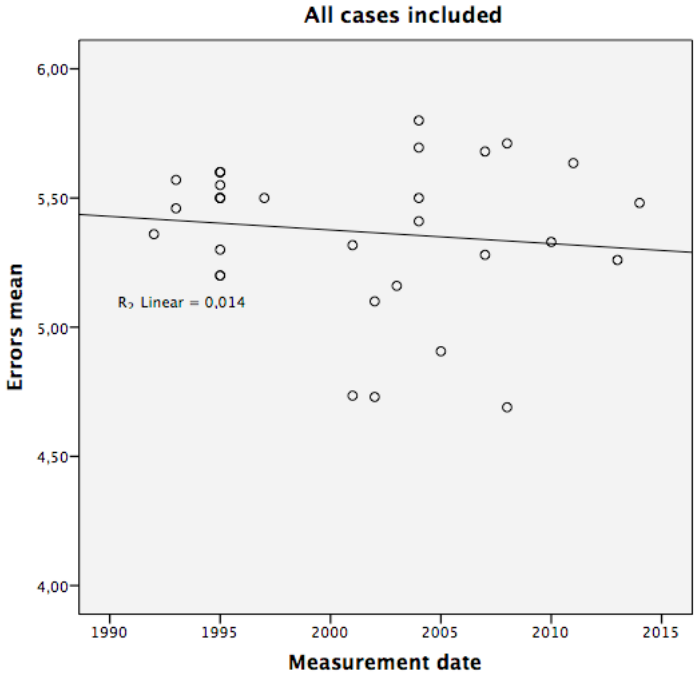


Figure 3.4: Here you can see the course of the Errors dimension over time. The concerns for this dimension seemed to be somewhat constant over time.

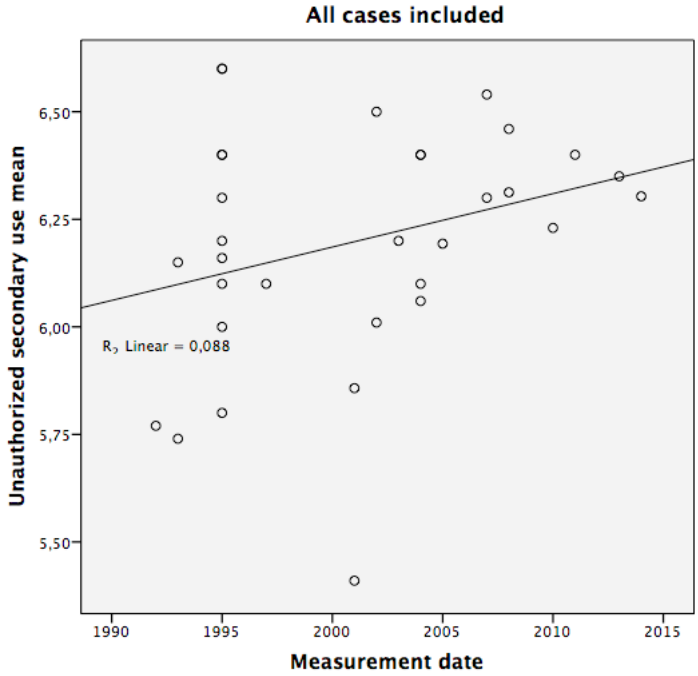


Figure 3.5: Here you can see the course of the secondary use dimension over time. The concerns for this dimension increased over time.

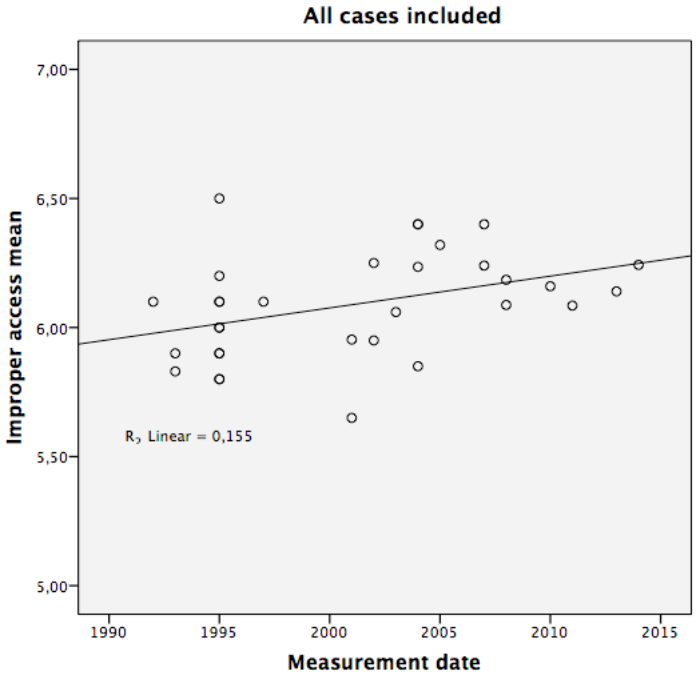


Figure 3.6: Here you can see the course of the Improper Access dimension over time. The concerns for this dimension increased over time.

Correlations

			CFIP mean	Collection mean	Unauthorized secondary use mean	Improper access mean	Errors mean
Spearman's rho	Measurement date	Correlation Coefficient	,186	-,079	,359*	,461**	-,061
		Sig. (2-tailed)	,284	,664	,044	,008	,740
		N	35	33	32	32	32

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 3.7: Here you can see the results of the bivariate correlation analysis of Concern For Information Privacy construct and its dimensions versus the dates when the respondents filled out the survey which measures their Concern For Information Privacy. Spearman's correlation coefficient was used for the analysis.

In Figure 3.7 you can find the magnitudes and significances of the correlations between the Concern For Information Privacy construct and the four dimensions, Collection, Unauthorized Secondary Use, Improper Access and Error, on the one hand, and the dates when the survey was administered on the other. Some data points did not specify certain dimensions.

As can be seen in Figure 3.7, the Collection and Errors dimensions did not yield a significant correlation with respect to time. However, the Unauthorized Secondary Use and Improper Access dimensions did have a significant positive correlation with respect to time. This means that over the past twenty years, in general people's concerns for Unauthorized Secondary Use and Improper Access of their personal information increased. These correlations had significances of $p_{sec} = 0.044$ and $p_{acc} = 0.010$. And these relations had correlation coefficients of $\rho_{sec} = 0.358$ and $\rho_{acc} = 0.449$. From this, the coefficients of determination could be derived: $R^2_{sec} = 0.128$ and $R^2_{acc} = 0.202$. So the change in concerns for Unauthorized Secondary Use and Improper Access was accounted for by the change in time by 12.8% and 20.2% respectively. Of course, it was not the simple change in time that has made people more concerned for the Unauthorized Secondary Use and Improper Access of their personal information. But other factors that have been changing over time may be responsible for this. These factors could include the changes in culture, the rise of the internet and other

technological developments.

> *It can be said with a certainty of respectively $p_{sec} = 0.044$ and $p_{acc} = 0.010$ that in general, people's concerns for Unauthorized Secondary Use and Improper Access of their personal information have increased over the past 20 years.*

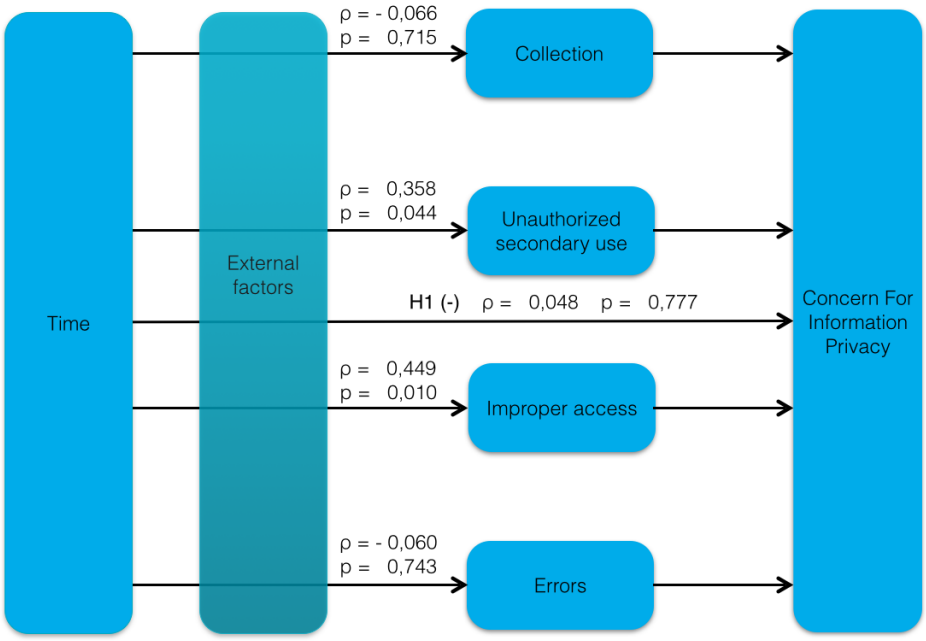


Figure 3.8: Here you can see the conceptual model of study 1 which was developed in the theoretical chapter, but now with the results from the correlation analyses integrated and also with indications which hypotheses were supported and which were not.

The results discussed above are presented in the conceptual model of Figure 3.8. The results of study 1 can easily be understood intuitively. Although insignificant, the correlation coefficient of both the Collection and Errors dimensions was negative. It can be speculated that technological development and the information technology revolution has caused people to get used to having their personal data being collected on a daily basis. This familiarization could have caused people to see the Collection of their personal data as less of a concern. This revolution also caused vast amounts of personal data to be recorded and used, which shifted the importance from data quality to data quantity. This could than in turn have caused people to be less concerned about the quality of their personal data, i.e. reduce the concerns for Errors. But of course, it should be emphasized that many of the observed correlations were in fact insignificant, rendering the discussion tentative.

3.2.2. SUPPLEMENTARY ANALYSES

In the methodology section it was discussed how applying certain selection criteria to the data set could increase the generalizability and reliability of the results. In this section, four supplementary analyses are presented of potentially interesting sub-selections of the aforementioned dataset. Again, the three selection criteria which were used are:

- Geographical location = United States
- Population type = field or students
- Context = organisational or online

1. STRICTEST SELECTION - THREE CRITERIA COMBINED (N = 9)

Applying all three of the above selection criteria left only nine data points up for analysis. Despite the low sample size, significant relations did arise. The computed correlation matrix can be found in Figure 3.9.

Correlations

			CFIP mean	Collection mean	Unauthorized secondary use mean	Improper access mean	Errors mean
Spearman's rho	Measurement date	Correlation Coefficient	,767*	,310	,881**	,619	,048
		Sig. (2-tailed)	,016	,417	,004	,102	,911
		N	9	9	8	8	8

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 3.9: Here you can see the results of the bivariate correlation analysis of Concern For Information Privacy construct and its dimensions versus the dates when the respondents filled out the survey which measured their Concern For Information Privacy. Spearman's correlation coefficient was used for the analysis. Only the data points after applying the three extra selection criteria were included.

Below you can find the scatterplots of the Concern For Information Privacy construct and its dimensions versus time in Figures 3.10 to 3.14.

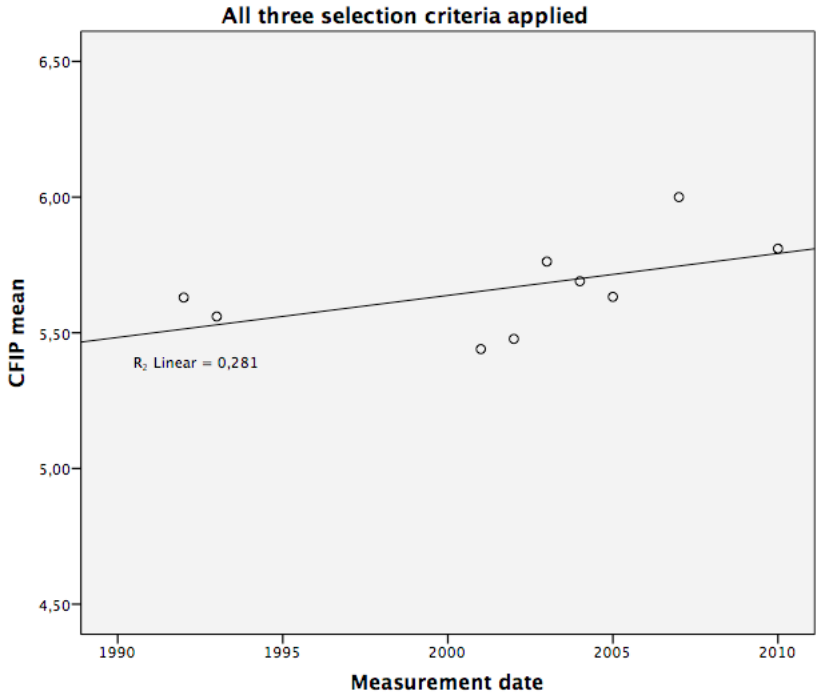


Figure 3.10: Here you can see the course of the Concerns For Information Privacy over time. The Concerns For Information Privacy increased significantly over time. Only the nine data points were included which were left after applying all three selection criteria.

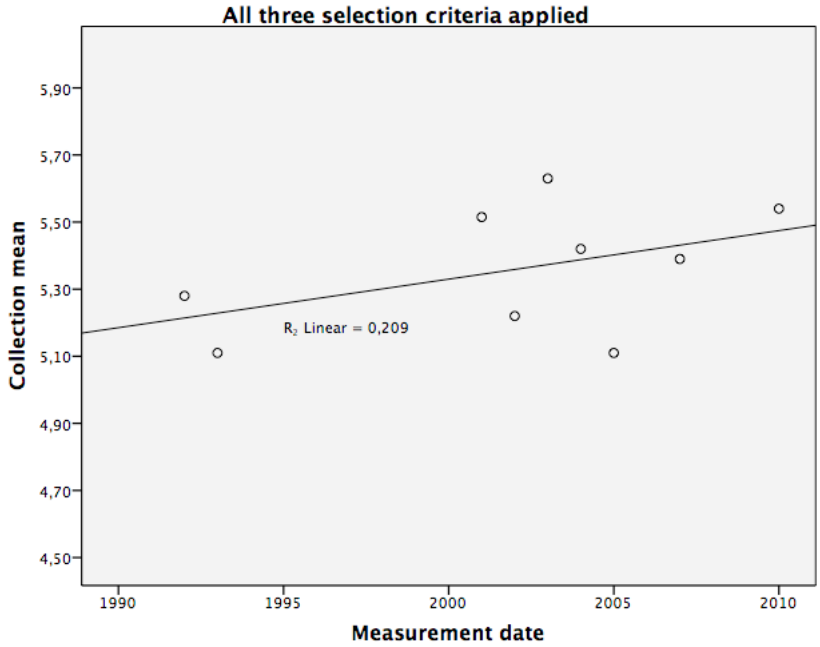


Figure 3.11: Here you can see the course of the concerns for Collection of personal data over time. These concerns seemed to increase over time but not significantly. Only the nine data points were included which were left after applying all three selection criteria.

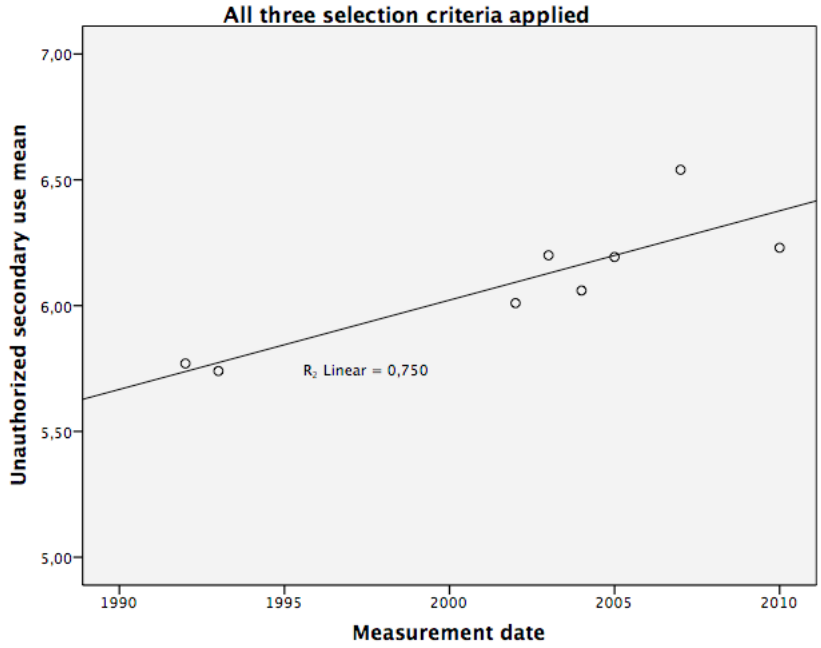


Figure 3.12: Here you can see the course of the concerns for Unauthorized Secondary Use over time. These concerns increased significantly over time. Only the eight data points were included which were left after applying all three selection criteria.

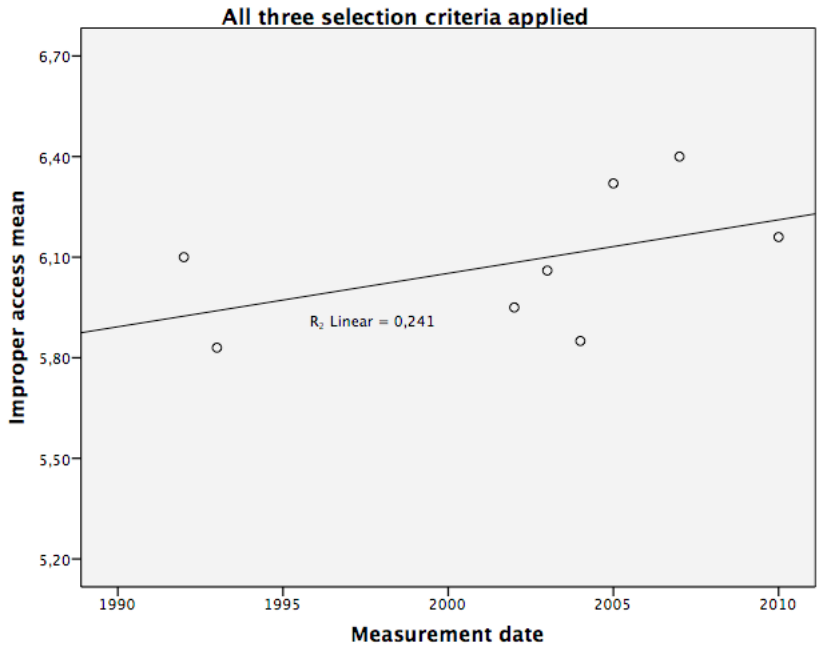


Figure 3.13: Here you can see the course of the concerns for Improper Access of personal data over time. These concerns seemed to increase over time but not significantly. Only the eight data points were included which were left after applying all three selection criteria.

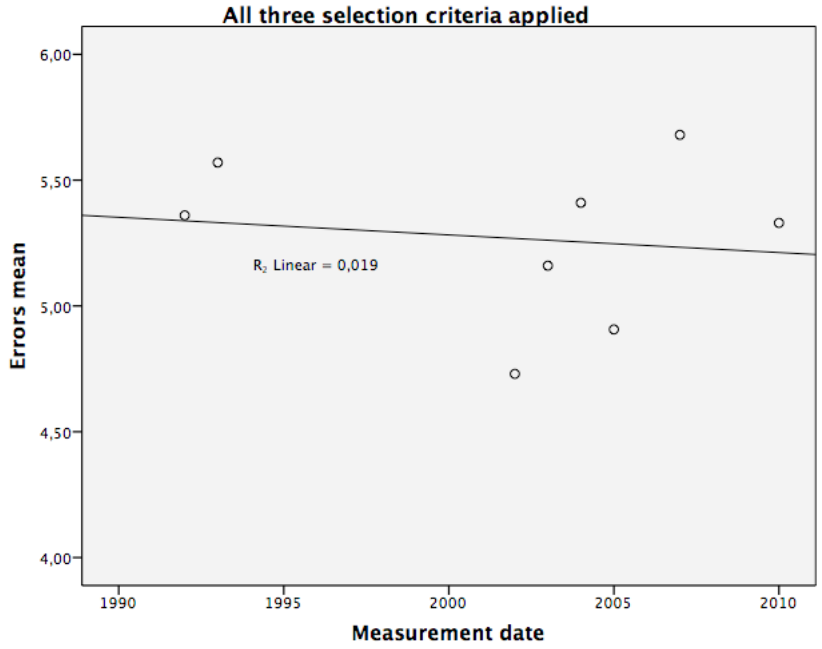


Figure 3.14: Here you can see the course of the concerns for Errors in personal data over time. These concerns seemed remain constant over time. Only the eight data points were included which were left after applying all three selection criteria.

In contrast to the results from the previous section, this analysis stated that Concerns For Information Privacy have increased over the past twenty years.

In the United States, in the field or student population, in an organisational or online context, Concerns For Information Privacy have increased over the past two decades.

The first hypothesis H1 was "Individual's overall Concerns For Information Privacy have changed over the past 20 years." It can now be concluded that in the United States, in the field or student population and in an organisational or online context that this hypothesis is approved. This is a massive result, because it was never quantitatively proven before that the Concerns For Information Privacy have actually increased.

Apart from this and also in contrast to the previous analysis, the insignificant correlation coefficients of the Collection and Errors dimensions versus time were positive, although still insignificant. Apparently the speculated explanation presented on page 34 did not hold for the sub-sample of Americans used in this analysis.

In Figure 3.15 you can find the conceptual model of study 1 with the results included based on the analyses after filtering by all three previously specified selection criteria. As discussed previously, these results were slightly different than the ones found in Figure 3.8, but were more specific and therefore more reliable.

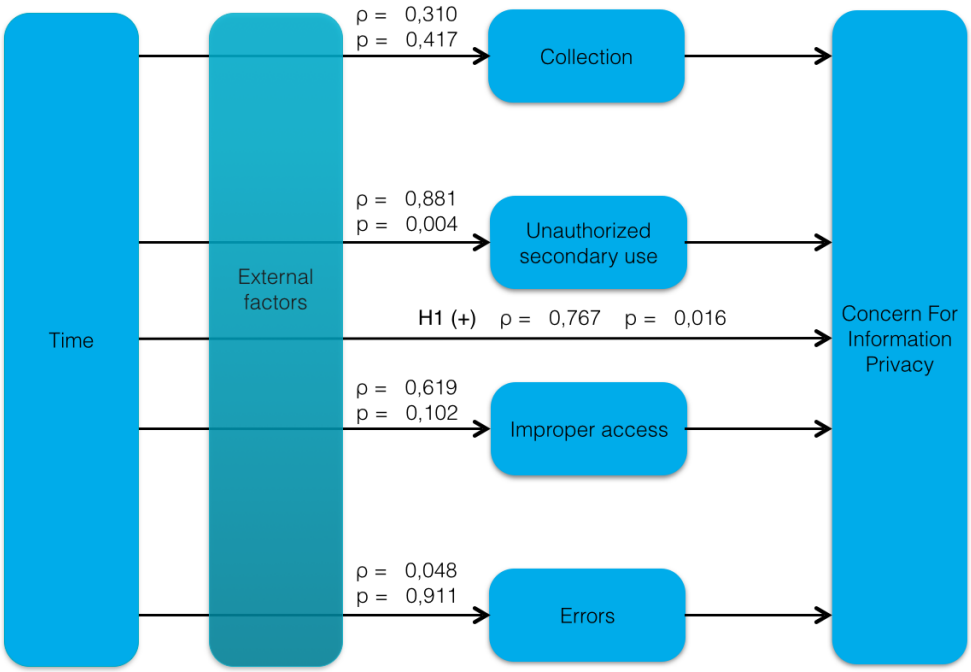


Figure 3.15: Here you can see the conceptual model of study 1 which was developed in the theoretical chapter, but with the results from the correlation analysis integrated and also with indications which hypotheses were supported and which were not. The results were based on the analysis after applying all three previously specified selection criteria.

It is now interesting to see which selection criteria created the discrepancy between the results of this analysis and the results of the analysis of the complete data set. This was investigated in the following sections by separately applying the three previously specified selection criteria.

2. GEOGRAPHICAL LOCATION AS A SINGLE SELECTION CRITERION (N = 18)

In the methodology chapter is was discussed that ideally the influence of other factors could be ruled out by keeping as much variables constant as possible. This was done in the previous section. To get a better grip of the data set and its versitility, these analyses explored the impact of applying the above discussed selection criteria separately instead of simultaneously. First, the geographical location was kept constant. Again, the country which was selected for this, is the United States. Simply because this country had enough data points to be able to find any correlations. The results for this bi-variate correlation analysis can be found in Figure 3.16.

Only cases from US included

Spearman's rho			CFIP mean	Collection mean	Unauthorized secondary use mean	Improper access mean	Errors mean
	Measurement date	Correlation Coefficient	,355	-,072	,739**	,594*	,001
		Sig. (2-tailed)	,148	,776	,001	,012	,996
		N	18	18	17	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 3.16: Here you can see the results of the bivariate correlation analysis of the Concern For Information Privacy construct and its dimensions versus the dates when the respondents filled out the survey which measured their Concern For Information Privacy. Spearman's correlation coefficient was used for the analysis and only the data points from the US were included.

The results in Figure 3.16 contain similar results compared to the previous analysis of all data points. In this analysis, there was again proof of the increase of concerns for Unauthorized Secondary Use and Improper Access of personal information over time. But from this analysis could be concluded that these increases were independent of the geographical location, since all included cases were from the United States. This can be said with certainty for the United States, but there is no reason to believe this does not hold for other countries, since the analysis of all data points included countries from all over the world and yielded similar results. Also, the correlation between time and the overall Concern For Information Privacy construct was insignificant.

3. RESEARCH CONTEXT AS A SINGLE SELECTION CRITERION (N = 27)

Identical to the section above, in this section the research context acted as the selection criterion. The contexts which were selected for this were the general organisational and online contexts because they were quite similar and together, they left a large enough sample size (N = 27). The results for this bi-variate correlation analysis can be found in Figure 3.17. Remarkably, the relation between the concerns for Unauthorized Secondary Use of personal information and time was insignificant. Similar to the analysis including all data points and the analysis with all three selection criteria applyied, the relation between the general CFIP construct and time was insignificant and the relation between the Improper Access dimension and time was significant.

Only cases in organizational or online context included

			CFIP mean	Collection mean	Unauthorized secondary use mean	Improper access mean	Errors mean
Spearman's rho	Measurement date	Correlation Coefficient	,339	,238	,329	,505*	-,248
		Sig. (2-tailed)	,083	,251	,116	,012	,243
		N	27	25	24	24	24

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 3.17: Here you can see the results of the bivariate correlation analysis of the Concern For Information Privacy construct and its dimensions versus the dates when the respondents filled out the survey which measured their Concern For Information Privacy. Spearman's correlation coefficient was used for the analysis and only the data points from an organisational or online context were included.

4. POPULATION TYPE AS A SINGLE SELECTION CRITERION (N = 21)

Identical to the two previous sections, in this section the population type acted as the selection criterion. The population types which were selected were student and field populations. This was done because these were similar, general populations and it was assumed that the privacy concerns of students resemble those of the general population. The "non-student" population type was also taken into the analysis, since this was also a general field population (but explicitly excluding students). The selection left around twenty data points up for analysis, depending on the dimension. The results for this bi-variate correlation analysis can be found in Figure 3.18.

Only cases with student of field population type included

			CFIP mean	Collection mean	Unauthorized secondary use mean	Improper access mean	Errors mean
Spearman's rho	Measurement date	Correlation Coefficient	,295	-,192	,535*	,339	-,033
		Sig. (2-tailed)	,194	,431	,022	,169	,897
		N	21	19	18	18	18

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 3.18: Here you can see the results of the bivariate correlation analysis of the Concern For Information Privacy construct and its dimensions versus the dates when the respondents filled out the survey which measured their Concern For Information Privacy. Spearman's correlation coefficient was used for the analysis and only the data points with a student or field population type were included.

The above analysis yielded only one significant relation, which was between time and the concern for Unauthorized Secondary Use. This relation was also observed in all previous analyses.

3.3. SHORT SUMMARY OF RESULTS

The objective of study 1 was to gain insight into the course of the Concerns For Information Privacy over the past 20 years. This was done by executing a longitudinal quantitative meta-analysis on the Concern For Information Privacy construct. This concluded that on a global scale, people's concerns for Unauthorized Secondary Use and Improper Access of their personal information have increased, and that for a general/student population in a general/online context the overall Concerns For Information Privacy have increased in the United States. These results imply that organisations should focus on decreasing people's Concerns For Information Privacy.

4

STUDY 2

4.1. METHODOLOGY

The goal of study 2 was to find out which factors are of influence to individual's Concerns For Information Privacy. To do this in a clear and convincing manner, it was chosen to conduct this research quantitatively. Therefore, a questionnaire was developed. Using a questionnaire allowed for the use of several analytical methods which could numerically prove the magnitude and significance of the relations between the Concerns For Information Privacy and other factors of interest.

4.1.1. RESEARCH OBJECTIVE

Where study 1 functioned to explore the development of the Concern For Information Privacy construct and its dimensions over time, study 2 investigated the influence of several new factors on the Concern For Information Privacy construct and its dimensions.

The objective of this study was to investigate the influence of several new and old factors on the Concern For Information Privacy construct and its dimensions.

For study 2 it was also chosen to work with the CFIP construct. This was done because, I would like to gauge people's general privacy concerns about organisational practices and this is exactly what the "Concern For Information Privacy" construct focuses on.

4.1.2. SAMPLE

To set up a generalizable and reliable research it was important to obtain a representative sample of the required population. With this in mind, I designed and distributed a custom-made survey via edX. A clear introduction of this Massive Open Online Courseware platform is provided in the following subsection.

4.1.3. EDX INTRODUCTION

EdX is a huge international platform for online learning on which the best universities all around the world make their knowledge available for free to anyone with an internet connection. Universities such as Harvard, MIT and Delft University of Technology share their knowledge via courses which are called Massive Open Online Courseware, or MOOC for short. People can follow these MOOC's at a predetermined pace or at their own pace to eventually, if successful, receive a personalized certificate which can be checked online to

prove its authenticity. The certificates of the edX platform are being acknowledged by many companies and organisations, making the education valuable and useful.

4.1.4. SURVEY ADMINISTRATION PROCEDURE

I developed and integrated my questionnaire into one of the MOOC's from Delft, namely the course AE 1110x "Introduction to Aeronautical Engineering". An example of how the integration looked can be found in Figure B.1 in appendix B. The survey was administered as follows:

In "week 0", prior to the start of the course, students could already access the MOOC to watch an introductory video, take a look at the course planning or prepare the upcoming lectures. On the Wednesday of week 0, the questionnaire was integrated into the edX platform, so that MOOC students could voluntarily fill out the survey.

4.1.5. QUESTIONNAIRE

The developed questionnaire was composed of existing items as well as newly developed items. This section presents a list of all the constructs which were measured by the questionnaire. Below, it is outlined for each construct why it was integrated into the questionnaire and how it was measured. A copy of the original questionnaire can be found in appendix A.

CONCERN FOR INFORMATION PRIVACY

The CFIP construct and its four dimensions were fully taken from existing literature (Smith *et al.*, 1996). As in study 1, the CFIP construct consists of the four dimensions Collection, Unauthorized Secondary Use, Improper Access and Errors. The complete measurement tool (all questions) was distributed among the participants. Please consult appendix A for the exact list of questions.

KNOWLEDGE OF PERSONAL DATA COLLECTION AND USAGE

It was fairly difficult to develop a scale which correctly measured an individual's Knowledge Of Personal Data Collection And Usage. Intuitively, the best way to do this was to ask respondents substantive questions and use the amount of right answers as a measure for the level of Knowledge Of Personal Data Collection And Usage. This could include asking for certain numbers or percentages which have to do with the subject and using the deviation as a measure for the level of knowledge. The selected way of measuring this concept was however to explicitly ask to how the respondents perceived their knowledge to be on the subject. An example statement is "I feel that I know enough about online personal data collection and usage to safely use services which require the collection of my personal data."

IMPORTANCE OF REALISTIC ALTERNATIVES

Nowadays, people are often forced into using services which require the collection of personal data. Online search engines, internet browsers, online shopping sites, texting services, email services, telephone services and social network sites have all crept into our daily lives and there is often no escaping from these services even if you would want to. That being said, is it not strange that these inescapable services collect and use so much of our personal data?

It was expected that people who agreed with the above statement and see the Importance Of Realistic Alternatives had more Concerns For Information Privacy, especially more concerns for the Collection of their personal data. The selected way of measuring this concept was by using the following single-item scale "I feel annoyed that I often have to use services which I rather would not, because of privacy reasons."

FREQUENCY OF USE OF SERVICES WHICH REQUIRE THE COLLECTION OF PERSONAL INFORMATION

The relation between the above construct "Importance Of Realistic Alternatives" and the Concern For Information Privacy was expected to depend on the Frequency Of Use Of Services Which Require The Collection Of

Personal Information. Therefore, this construct was also measured in the questionnaire and tested as a mediating variable. This construct was measured using the following single item construct "I often use online services which require the collection of my personal data."

PERSONAL DEVELOPMENT OF CONCERNS FOR INFORMATION PRIVACY

In study 1 it was investigated whether the Concern For Information Privacy of society in general had changed over time. To check this result, two questions were integrated into the questionnaire which measured individuals' long term personal perception of the change of their Concerns For Information Privacy. Although this is very difficult to self-report these kind of questions objectively, it was very interesting to see how people thought their perception of information privacy had shifted over time. It was also interesting to see whether the result of these questions would or would not be in line with the findings of study 1. An example question of this construct is "Over the past ten years my attitude towards information privacy has become more and more tolerant."

WILLINGNESS TO DISCLOSE IN A MOOC SETTING

When measuring privacy perceptions in a self-reporting manner, the privacy paradox should always be thought of. This privacy paradox is the discrepancy between people's intentions in data disclosure and their actual behaviour when disclosing data. It has been proven that people would rather not share as much data, but when push comes to shove they disclose it without much difficulty (NORBERG *et al.*, 2007). This peculiar relation between the Willingness To Disclose and CFIP could be different in a MOOC setting. Therefore, a single-item scale was added to measure the Willingness To Disclose In A MOOC Setting: "I would you be happy if edX researchers would use any additional available data of me to improve the quality of this research." Note that the question was phrased in a peculiar way so that behaviour was measured instead of intention.

UNIVERSAL HUMAN VALUES (SCHWARTZ, 1994)

In the current information privacy literature a lot of research explored the possible antecedents of the Concern For Information Privacy. Research explored many antecedents of CFIP, such as morality, self efficacy, risk taking, trust, anxiety (Korzaan *et al.*, 2009), perceived vulnerability, perceived ability to control (Dinev and Hart, 2004), perceived severity, response efficacy and reward (Mohamed and Ahmad, 2012). Quite remarkably, general universal human values were never considered as possible antecedents for the CFIP construct. That was why this was tested in this research. Unfortunately, due to a lack of space in the survey, it was only possible to integrate five from a total of ten universal human values as defined by Schwartz (1994). The five values of which the largest correlation with the CFIP was expected were included. These were Universalism, Self-Direction, Stimulation, Hedonism and Security.

The lack of space in the questionnaire also forced me to merge multiple items of the above stated values into single item scales. This was done for Universalism, Stimulation, Hedonism and Security. Although the used questions still greatly resembled the original developed portrait value questions, this could have affected the results. And this also prevented the execution of Cronbach's internal reliability test. Only the items for the human value construct Self-Direction was kept in the original form. To consult the exact questions please see appendix A.

The above could have affected the reliability of the outcomes, so the results should only be seen as an indication of the relation between the Concern For Information Privacy and the measured universal human values.

INTERNET ACTIVITY

An individual's internet activities and online behaviour could have also had an impact and fuel or dissolve ones Concerns For Information Privacy. Because of this, the following constructs were integrated into the questionnaire:

- Frequency Of Social Media Use;
- Frequency Of Online Forum Use;
- Average Time Spend Online Daily.

It was expected that the Frequency Of Online Forum Use, Average Time Spend Online Daily and the Frequency Of Social Media Use were negatively correlated with the CFIP construct and its dimensions.

This was because the Frequency Of Online Forum Use and the Average Time Spend Online Daily confront people more and more with the sharing of personal data. This confrontation was expected to cause familiarisation and dissolve people's Concerns For Information Privacy. These constructs were measured with the single-item scales "How often do you contribute questions or answers to online forums?" and "How much time do you averagely spend online in a day?"

The Frequency Of Social Media Use was a measure for the voluntary disclosure of personal data. It was expected that people who often share their personal data voluntarily were less concerned about their personal data. This construct iwas measured with the single item scale "How often do you use social media (i.e. Facebook, Twitter, Google+, Weibo, etc.)?"

A problem was encountered for the "Average Time Spend Online Daily" construct when the questionnaire data was received. Even though the questionnaire clearly stated that the respondent is asked to fill out their Average Time Spend Online Daily in *minutes*, often the respondents mistakenly gave their answer in *hours*. This could be said with certainty because participation in the Aeronautical Engineering course, in which this survey was integrated, took at least 3,5 hours per week, which was a minimum of 30 minutes per day. To correct these mistakes, a wide margin was taken and the responses of "14" or less were multiplied by 60 to represent the correct amount of time in minutes.

FINANCIAL WEALTH

To find out if there were any correlations between income, Financial Stress and Financial Stability and the Concern For Information Privacy, the following constructs were integrated into the questionnaire:

- Financial Stability;
- Financial Stress;
- Financial Independence;
- Household Size;
- Household Income.

The construct Financial Independence was of great importance, since people who were not financially independent could disrupt the data very easily since they rarely felt Financial Stress and were almost always financially stable. Looking at Financial Stability, Financial Stress and Financial Independence allowed for a more thorough exploration of the relation between wealth and the CFIP.

The constructs Household Size and Household Income together formed an estimate for the amount of wealth of an individual, by calculating the Household Income Per Household Member. These questions were only posed to respondents from the United States and Canada, because this required familiarisation with dollars. For example questions of the above constructs please consult appendix A.

DEMOGRAPHICS

The followings demographics were integrated into the questionnaire:

- Gender;
- Age;
- Marital Status;
- Highest Level Of Education;
- Position In Society;
- Occupation;
- Nationality;
- Current Residence;
- Latitude And Longitude Information.

The main reason for integrating these demographics into the questionnaire was to assess the diversity of the sample population. But it also allowed to find many interesting correlations. For instance, Age, Highest Level Of Education, Position In Society and Occupation were all potential estimators for the Concern For Information Privacy construct and its dimensions. The Latitude And Longitude Information allowed for the execution of a geographical cluster analysis in which differences between geographical clusters were explored. For the complete list of demographic questions please consult appendix A.

4.1.6. RELIABILITY OF NEW SCALES

To attain a high reliability for all the above measured constructs, it was best to solely use old scales which were already tested thoroughly in literature. Therefore old pre-tested scales were often used. It was for example very important that it was certain that the Concern For Information Privacy construct and its dimensions Collection, Unauthorized Secondary Use, Improper Access and Errors were truly reliable, since the entire study was based around these constructs. Despite the above, it was sometimes necessary to create new scales. This was the case for Personal Development Of Concerns For Information Privacy, the level of Knowledge Of Personal Data Collection And Usage, the Importance Of Realistic Alternatives and the Frequency Of Use Of Services Which Require The Collection Of Personal Information. These scales were newly created in this study since they were not used before in literature. Normally, before introducing a new scale into the research field, elaborate testing should be executed to confirm the validity, generalizability and reliability of a new scale. However, considering the time span of this thesis it was decided to exclude this from the scope.

4.1.7. DATA PREPARATION AND SAMPLE SIZE

OUTLIERS

Several unrealistic outliers were discovered and deleted from the data set. For example, individuals which specified to be younger than 12 years old or older than 85 years old were excluded from this research. This were 7 individuals. Also, people who indicated they had more than 50 people living in their household were deleted from the research, since this seemed unrealistic.

A classical outlier analysis indicated that the Concern For Information Privacy construct included several outliers on the "left" or "unconcerned" side. However, this construct was heavily left-skewed, which could have affected this outlier analysis. Also, there was no objective reason to remove these outliers from the data set, since some people were genuinely not concerned about their information privacy. Therefore these data points were kept in the data set.

FILTERING PROCESS

In the filtering process all respondents of which a single or several essential variables were missing were completely removed from the research. This excluded the financial related questions, such as Financial Independence, Financial Stress, Financial Stability and Household Income, because these values were often missing. This was caused due to the fact that these questions were only posed to respondents from the US and Canada. Apart from this, a very large amount of respondents just completed a small section of the questionnaire. This was because the integration of the questionnaire into the MOOC made the survey very easily accessible. This caused curious but unmotivated respondents to start the survey, without ever completing it.

The pre-filter data set contained 2036 responses and after this initial filtering 838 entries were left. An additional filtering method based on location accuracy was also applied to improve the reliability of the sample (see next section). This filtered 40 respondents, which led to the final sample size $N = 798$.

Since this method did not require to send the questionnaire to people, a conventional response rate could not be calculated. The amount of students who enrolled for the "Introduction to Aeronautical Engineering" MOOC was 7151, but this number should not be seen as the initial sample population. A more representative number was the amount of active students, or students which had logged in at the beginning of the course. This were 3758 students. This resulted in a gross response rate of 54.2 % and a net response rate of 21.2 %.

LOCATION ACCURACY CHECK

After the previously discussed initial filtering, a unique filtering method was used to exclude unreliable respondents. In the questionnaire people were asked to indicate in which country they currently resided. However, the Latitude And Longitude Information of their current position was also extracted from their IP address. This allowed to perform a location accuracy check which was used as an indication of the reliability of the respondent. Moreover, the respondents which answered the Current Residence question unreliably were excluded from the research. This analysis was done in the program R and the code which was used for this analysis can be found in appendix C in Figure C.1 (R Development Core Team, 2008).

The above location accuracy check was unfortunately susceptible to errors. The acquired longitude and latitude coordinates were extracted from the respondents' IP addresses. This extraction was computed by the questionnaire software which was used, called Qualtrics. Qualtrics states that

"... the location is an approximation determined by comparing the respondent's IP address to a location database. Inside the United States, this data is typically accurate to the city level. Outside the United States, this data is typically only accurate to the country level." (Qualtrics, 2014).

So the accuracy of the Latitude And Longitude Information on the country level should suffice, but an IP address was not a 100 percent reliable source for ones geographical location because it could be tampered with. If for example a respondent would have filled out the survey through a proxy server, which could be done very easily through sites like www.hidemyass.com, the location accuracy check would flag this respondent as unreliable. It happens quite often in practice that governments censure certain parts of the internet and consequently people simply use a proxy server to work around this censure. Countries which perform internet censorship are China, Cuba, Iran, North Korea, Saoedi Arabia, Syria, Tunisia and many others.

However, the location accuracy check flagged only 40 respondents, which was very little compared to the total amount of respondents. So even though it was a waste to delete these data points, since some of them were probably clean data, it could be concluded that the remaining 798 data points were more reliable, which was very much worth the loss of 5 % of the data.

DID THE LOCATION ACCURACY CHECK AFFECT THE DATA?

During the development of this thesis, it was criticized that the above described location accuracy check affected the data set and therefore affected the results. The line of argumentation was that people who had a very specific opinion on the topic of information privacy were more likely to have tampered with their IP address or were more likely to lie about their location for privacy reasons. Therefore the group of respondents filtered by the above described procedure would be a very critical and important group.

To refute this criticism a statistical test was executed to find out whether there were significant differences between the filtered respondents and the remainder of the dataset. First of all a χ^2 test was executed between the CFIP and the "Result of location accuracy check" variables. For this test to be reliable, certain assumptions had to be met. In the cross-tabulation the expected count of a single cell could not be lower than one and the percentage of cells with an expected count lower than five could not be below 20 %. To abide these assumptions, the CFIP variable was split up into four categories based on the quartiles of the entire data set: low concern, below average concern, above average concern and high concern. The χ^2 test between these defined groups of the Concern For Information Privacy and the two groups of location accurate and location inaccurate respondents can be found in Figure 4.1 and the corresponding crosstabulation can be found in Figure 4.2. The results showed that assumptions of the χ^2 test were not violated and group filtered by the location accuracy check did not differ significantly from the other respondents ($p = 0.626$)

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1,750[a]	3	,626
Likelihood Ratio	1,808	3	,613
Linear-by-Linear Association	,310	1	,578
N of Valid Cases	838		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 9,50.

Figure 4.1: Here you can see results of the χ^2 test between the four groups of CFIP based on the quartiles of the data set and the two groups of location accurate and location inaccurate responses.

Result of location accuracy check * CFIP categorized based on quartiles Crosstabulation

			CFIP categorized based on quartiles				
			Low concern	Below average concern	Above average concern	High concern	Total
Result of location accuracy check	Accurate	Count	202	204	203	189	798
		Expected Count	201,9	200,9	205,7	189,5	798,0
	Not accurate	Count	10	7	13	10	40
		Expected Count	10,1	10,1	10,3	9,5	40,0
Total		Count	212	211	216	199	838
		Expected Count	212,0	211,0	216,0	199,0	838,0

Figure 4.2: Here you can see crosstabulation matrix with the counts and expected counts of the four groups of CFIP based on the quartiles of the data set versus the two groups of location accurate and location inaccurate responses.

The results in Figure 4.1 disproved the previously presented criticism, but could be explained by the fact that respondents who used a proxy server to tamper with their IP adress or lied about their Current Residence, felt less concerned for their information privacy *because* they had taken these measures. This however was out of the scope of this study. Either way, the χ^2 test concluded that the group filtered by the location accuracy check did not differ significantly from the remainder of the data set. So these respondents were filtered from the analyses in the following section to improve reliability.

4.2. RESULTS

4.2.1. DEMOGRAPHICS

In Table 4.1 you can find the demographics of the sample in this study. Several things stood out here, the most significant perhaps being the large percentage of male respondents. (86.8%) This was caused by the technical character of the Aeronautical Engineering topic of the MOOC in which the survey was integrated. Apart from this the demographics were nicely spread, with small augmentation to lower Age, higher education and a student occupation.

Profile	Items	Frequency	Percentage
Gender	Male	693	86.8
	Female	101	12.7
	Other	4	0.5
Age	< 20	133	16.7
	20 - 29	422	52.9
	30 - 39	118	14.8
	40 - 49	67	8.4
	50 - 59	31	3.9
	> 59	27	3.4
Education level	Doctorate	31	3.9
	Master's degree	147	18.4
	Bachelor's degree	289	36.2
	Associate's degree	53	6.6
	High school	246	30.8
	Less than high school	16	2.0
	Other	16	2.0
Ethnicity	Caucasion / White	288	36.1
	Asian Indian	178	22.3
	Hispanic / Latino	115	14.4
	Other Asian	75	9.4
	African American / Black	28	3.5
	Arab	22	2.8
	Native Indian	22	2.8
	Other	70	8.8
Occupation	Student	391	49.0
	Professional	304	38.1
	Unemployed	62	7.8
	Other	41	5.1

Table 4.1: Here you can find the demographics of the sample.

4.2.2. INTERNAL VALIDITY

Construct	Cronbach's Alpha	No. of Items
Collection	0.820	4
Errors	0.842	4
Secondary Use	0.824	4
Improper Access	0.829	3
Knowledge Of Collection And Usage	0.819	2
Personal Development Of CFIP	0.492	2
Self-Direction	0.633	2

Table 4.2: List of the multi-item scales used in study 2 and their corresponding Cronbach alpha values.

In Table 4.2 it can be seen that the constructs used from the questionnaire of Smith, Milberg, and Burke all had sufficient internal reliability when the values of Cronbach's Alpha were compared with the conventional minimum of $\alpha = 0.7$. The "Personal Development Of CFIP" construct however turned out to have a very weak internal validity. Therefore, for the remainder of this thesis, I used the item "Over the past ten years my attitude towards information privacy has become more and more tolerant" as a single item scale to estimate people's personal long-term perception of the change in his/her Concerns For Information Privacy.

Four of the five universal human values used in this research were converted to single item scales to save space in the questionnaire. The value Self-Direction however was conserved in its original 2-item form, which allowed for the computation of Cronbach's alpha. Despite the original form, the two items resulted in an internal reliability of only $\alpha = 0.633$, which was below the conventional minimum. However, literature also often assesses values of Cronbach's alpha of $0.6 < \alpha < 0.7$ as "Acceptable" (George and Mallery, 2003). Because of this and the fact that I used the original previously tested portait value questions of Schwartz, the internal validity of the Self-Direction construct was considered to be sufficient.

Structure Matrix

	Factor			
	1	2	3	4
ACCESS N	,839	,350	,533	,178
SEC M	,812	,361	,430	,344
ACCESS I	,771	,318	,559	,215
SEC K	,755	,332	,407	,324
ACCESS D	,703	,342	,516	,522
SEC G	,662	,327	,430	,312
COLL J	,374	,882	,302	,084
COLL O	,392	,735	,246	,074
COLL A	,175	,682	,147	,189
COLL E	,415	,647	,274	,271
ERRORS L	,473	,246	,850	,066
ERRORS H	,538	,262	,799	,083
ERRORS F	,456	,189	,767	,203
ERRORS B	,355	,230	,626	,217
SEC C	,607	,338	,400	,738

Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization.

Figure 4.3: Here you can see the factor matrix which resulted from the confirmatory factor analysis of the CFIP construct. The Oblimin rotation method and maximum likelihood extraction method were used.

In Figure 4.3 the result of a confirmatory factor analysis of the CFIP construct can be found. This analysis was done to verify whether this measurement tool correctly measured the corresponding four dimensions of the Concern For Information Privacy construct. The result however was very unexpected. It showed that the Collection and Errors dimensions were clear factors with convincing factor loadings. But the Unauthorized Secondary Use and Improper Access dimensions however had merged into one and the same factor, leaving only three clearly distinguishable factors.

This could be explained by the obsolescence of the questionnaire, since the questionnaire was twenty years old. But this line of thought was excluded by the fact that this same clustering was found much earlier in a research by Campbell (1997).

Another reason could be that the sample population had caused the merger of these two dimensions. Apparently, the international population interested in Aeronautical Engineering did not distinguish between Unauthorized Secondary Use and Improper Access. Although these dimensions were definitely not the same, they did resemble one another. For example, the question: "Computer databases that contain personal information should be protected from unauthorized access - no matter how much it costs." was an item for the Improper Access dimension, since unauthorized access was considered improper.

It could be argued that people nowadays worry about Improper Access in general to their personal information, whether this is Unauthorized Secondary Use or some other improper way of accessing their personal data. Because if people's personal information is used for secondary goals without authorization, they do not care about the actual use of their information, but about their personal information being accessed improperly.

4.2.3. GEOGRAPHICAL CLUSTER ANALYSIS

In Figures 4.4 and 4.5 you can find the latitude and longitude coordinates of the created cluster centers and the amount of cases in each cluster respectively. The cluster numbers were defined as follows:

1. Africa
2. North America
3. Europe
4. South America
5. Oceania
6. Asia

In Figure 4.6 you can find the same defined cluster centres as in Figure 4.4, but than visualized on the world map.

Final Cluster Centers

	Cluster					
	1	2	3	4	5	6
Location Latitude	-,78850600	34,9667857	46,0850755	-12,159349	-33,217160	19,9376034
Location Longitude	14,8374095	-91,334607	10,2846214	-58,275315	145,059288	83,4992990

Figure 4.4: Here you can see the locations of the centers of the six geographical clusters indicated by latitude and longitude coordinates.

Number of Cases in each Cluster

Cluster	1	17,000
	2	155,000
	3	242,000
	4	106,000
	5	15,000
	6	263,000
Valid		798,000
Missing		,000

Figure 4.5: Here you can see the amount of cases included in the six geographical clusters.

Test Statistics[a,b]

	Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Chi-Square	10,144	55,461	24,473	3,976	6,227
df	5	5	5	5	5
Asymp. Sig.	,071	,000	,000	,553	,285

a. Kruskal Wallis Test

b. Grouping Variable: Geographical cluster Number

Figure 4.7: Here you can see the results of the independent t test of the Concern For Information Privacy construct and its dimensions for the different geographical clusters.

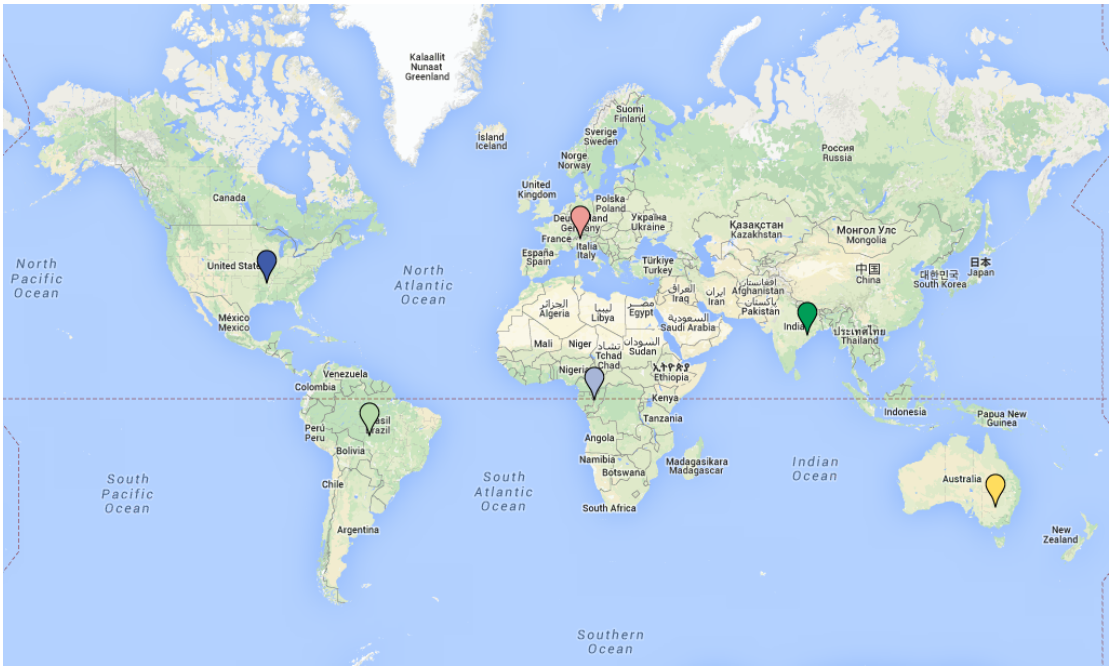


Figure 4.6: Here you can see the defined cluster centers visualized on the world map.

To find out whether there were significant differences in the Concern For Information Privacy and the four dimensions between the clusters, a Kruskal Wallis Chi-Squared test was executed. The results of this test can be found in Figure 4.7. Quite surprisingly the general Concern For Information Privacy construct did not differ significantly between the clusters. The Collection and Errors dimensions however did have significant differences between clusters. These differences had a magnitude of $\chi^2 = 55.461$ and $\chi^2 = 24.473$ respectively, both with a significance of $p = 0.000$.

> *Apparently the concerns for the Collection of personal data and the concerns for Errors in personal data vary greatly across continents and across cultures, whereas the concerns for Unauthorized Secondary Use and Improper Access of personal data are more similar across continents and cultures.*

Previous literature showed that the general Concerns For Information Privacy do vary with geographical location (Bellman *et al.*, 2004; Milberg *et al.*, 1995) Accordingly, hypothesis H12 which was developed for this topic was "Levels of Concerns For Information Privacy will differ across countries." Despite several dimensions significantly varying across countries, this hypothesis was rejected because the overall Concern For Information Privacy construct did not yield a significant result.

Perhaps the result of Milberg *et al.* (1995) could not be replicated because of the higher than average educated sample with a high percentage of male respondents. Due to this insignificant result the geographical cluster analysis was not elaborated on any further.

4.2.4. WEALTH AFFECTING CFIP

In Figure 4.8 you can find the correlation matrix between Household Income, Household Income Per Household Member and CFIP and the dimensions, which showed that people with a higher income per household member had more Concerns For Information Privacy. To further investigate this relation, another analysis was executed with only financially independent individuals in the sample. This analysis can be found in Figure 4.9. Quite surprisingly not the "Household Income Per Household Member", but Household Income in general was found to have a positive correlation with the CFIP and the Unauthorized Secondary Use and Improper Access dimensions. Although slightly different both of these results indicated a positive relation

between income and the CFIP.

> *So it was concluded that there is a positive relation between income and the Concern For Information Privacy.*

The hypothesis H5C was "Individuals with a higher Household Income Per Household Member are more concerned for their information privacy." This hypothesis was approved because of the above result. The other hypothesis on income was H5E: "Individuals with a higher Household Income are more concerned for their information privacy". This hypothesis was approved only for financially independent individuals. (see below)

Correlations

			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Household income	Correlation Coefficient	,106	,030	,034	,146	,129
		Sig. (2-tailed)	,316	,777	,749	,168	,225
		N	91	91	91	91	91
	Income per household member	Correlation Coefficient	,216*	,130	,158	,260*	,185
		Sig. (2-tailed)	,042	,224	,140	,014	,082
		N	89	89	89	89	89

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.8: Here you can see the results of the bivariate correlation analysis of the Household Income and Household Income Per Household Member and the CFIP construct.

Financial independency = Yes

Correlations[a]

			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Household income	Correlation Coefficient	,250*	,127	,149	,258*	,249*
		Sig. (2-tailed)	,047	,318	,240	,039	,047
		N	64	64	64	64	64
	Income per household member	Correlation Coefficient	,203	,118	,134	,210	,148
		Sig. (2-tailed)	,114	,359	,298	,101	,251
		N	62	62	62	62	62

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Financial independency = Yes

Figure 4.9: Here you can see the results of the bivariate correlation analysis of the Household Income and Household Income Per Household Member and the CFIP construct. This analysis was identical to the one presented in Figure 4.8, but here only financially independent individuals were included into the analysis.

To further investigate the influence of wealth on the Concern For Information Privacy, I looked at the effect of Financial Stability and Financial Stress. In Figure 4.10 you can find the spearman's correlation matrix of Financial Stability and Financial Stress versus the CFIP construct and its dimensions. In the matrix the distinction was made between financially dependent and independent individuals. Thus, for financially dependent people there was no significant relation between CFIP and Financial Stability or Financial Stress. Financially stable people however tended to have more Concerns For Information Privacy, with three of the four dimensions having a significant positive correlation. Another small result was that people with more Financial Stress were more concerned about Errors in their personal data.

Correlations

Financial independence				Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Yes	Financial stability	Correlation Coefficient	,150**	,111*	,081	,141**	,112*
			Sig. (2-tailed)	,003	,030	,112	,006	,028
			N	384	384	384	384	384
		Financial stress	Correlation Coefficient	,075	,036	,120*	,009	,047
			Sig. (2-tailed)	,143	,488	,019	,856	,356
			N	384	384	384	384	384
	No	Financial stability	Correlation Coefficient	-,015	-,058	,006	,024	,023
			Sig. (2-tailed)	,757	,241	,909	,633	,640
			N	414	414	414	414	414
		Financial stress	Correlation Coefficient	,028	,094	,024	-,058	-,023
			Sig. (2-tailed)	,566	,056	,621	,237	,634
			N	414	414	414	414	414

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.10: Here you can see the results of the bivariate correlation analysis of the Financial Stability and Financial Stress variables versus the Concern For Information Privacy construct and its dimensions. A grouping was applied for financial dependent and financial independent respondents, since Financial Independence greatly impacted the relations presented in the correlation matrix.

The hypotheses H10A and H10B developed for these relations were "Financially independent individuals with more Financial Stress will have more Concerns For Information Privacy" and "Financially independent individuals with more Financial Stability will have more Concerns For Information Privacy," respectively. Because of the above stated results, H10A was rejected and H10B was approved.

In Figure 4.11 you can find the descriptives of the CFIP and its dimensions for financially dependent and independent individuals. In Figure 4.12 you can find the independent samples t test which statistically tested the significance of the difference between these groups. It turned out that financially independent people had more Concerns For Information Privacy than financially dependent people, especially more concerns for the Collection and Unauthorized Secondary Use of their personal information. This was anticipated because of the influence of the underlying factor Age. (Financially independent individuals are usually older) The hypothesis H10C developed to test this relation was: "Individuals who are financially independent will have more Concerns For Information Privacy than individuals who are financially dependent". The previously discussed results confirmed this hypothesis.

Group Statistics

	Financial independency	N	Mean	Std. Deviation	Std. Error Mean
Concern For Information Privacy	Yes	384	5,9123	,73380	,03745
	No	414	5,7451	,84291	,04143
Collection dimension	Yes	384	5,3835	1,10612	,05645
	No	414	4,9076	1,22416	,06016
Errors dimension	Yes	384	5,4974	1,06342	,05427
	No	414	5,5344	1,01137	,04971
Unauthorized secondary use dimension	Yes	384	6,4141	,84780	,04326
	No	414	6,2832	,99192	,04875
Improper access dimension	Yes	384	6,3542	,84212	,04297
	No	414	6,2552	1,01353	,04981

Figure 4.11: Here you can see the descriptives of the two groups of financially dependent and financially independent respondents for the Concern For Information Privacy construct and its dimensions.

Independent Samples Test

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Concern For Information Privacy	Equal variances assumed	2,978	796	,003	,16715	,05613	,05697	,27734
	Equal variances not assumed	2,993	792,850	,003	,16715	,05584	,05754	,27677
Collection dimension	Equal variances assumed	5,746	796	,000	,47585	,08281	,31330	,63841
	Equal variances not assumed	5,768	795,461	,000	,47585	,08250	,31392	,63779
Errors dimension	Equal variances assumed	-,504	796	,614	-,03702	,07345	-,18121	,10716
	Equal variances not assumed	-,503	783,686	,615	-,03702	,07359	-,18148	,10743
Unauthorized secondary use dimension	Equal variances assumed	1,996	796	,046	,13085	,06556	,00215	,25955
	Equal variances not assumed	2,008	790,787	,045	,13085	,06518	,00290	,25880
Improper access dimension	Equal variances assumed	1,493	796	,136	,09893	,06624	-,03110	,22897
	Equal variances not assumed	1,504	786,656	,133	,09893	,06579	-,03021	,22807

Figure 4.12: Here you can see the results of the independent samples t test between the two groups of financially dependent and financially independent respondents for the Concern For Information Privacy construct and its dimensions.

4.2.5. RELATION BETWEEN KNOWLEDGE AND CFIP

In Figure 4.13 you can find the correlations between Knowledge Of Personal Data Collection And Usage and the Concern For Information Privacy and its dimensions.

Correlations

			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Level of knowledge of personal data collection and usage	Correlation Coefficient	,135**	,044	,246**	,035	,091**
		Sig. (2-tailed)	,000	,218	,000	,319	,010
		N	798	798	798	798	798

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.13: Here you can see the results of the bivariate correlation analysis of the Knowledge Of Personal Data Collection And Usage versus the Concern For Information Privacy construct and its dimensions. Spearman's correlation coefficient was used.

In Figure 4.13 it can be seen that the Knowledge Of Personal Data Collection And Usage had small but significant correlations with the Concern For Information Privacy construct and also with two dimensions. At first glance one would conclude that the more people know about personal data collection and usage, the more they are concerned with their personal data. But when this relation was visualized using a scatterplot, which can be found in Figure 4.14, a different relation surfaced.

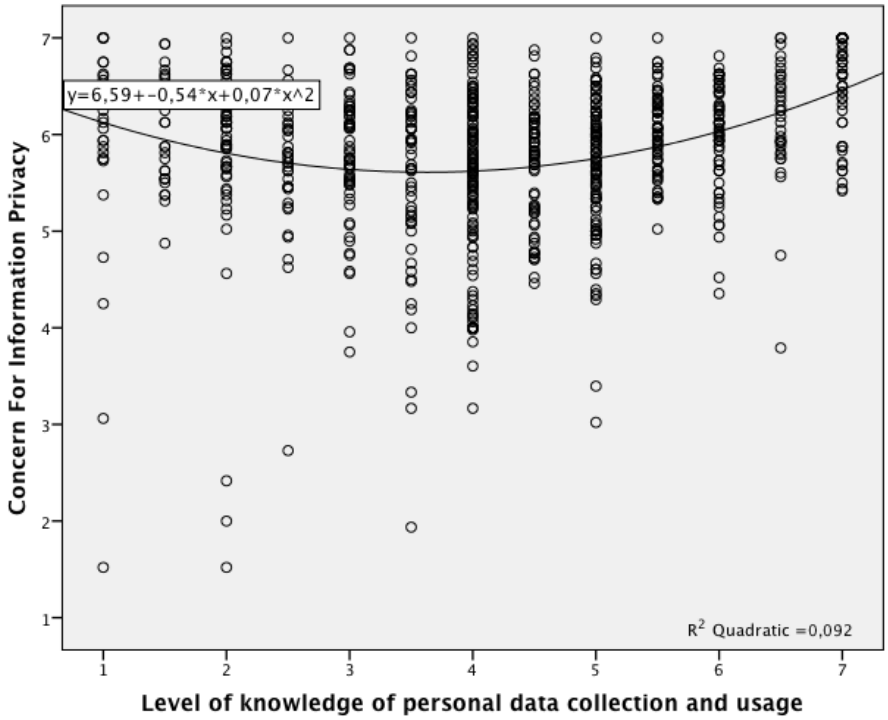


Figure 4.14: Here you can see a scatterplot of the degree of Knowledge Of Personal Data Collection And Usage versus the Concern For Information Privacy. Because of the apparent quadratic form, the scatterplot was fitted with a parabola.

In Figure 4.14 it can be seen that the degree of Knowledge Of Personal Data Collection And Usage had a non-monotonic parabolic relation with the Concern For Information Privacy. To test whether this is statistically significant, two bivariate correlation analyses were carried out on both sides of the cut-off point. The cut-off point was the point where the fitted parabola's slope equaled zero.

$$x_{cut-off} = \frac{0.54}{2 \cdot 0.07} = 3.86$$

So the first correlation analysis was executed on data points ranging from 1.0 to 3.5 and the second on data points ranging from 4.0 to 7.0. The results of these analyses can be found in Figure 4.15.

Correlations

	Level of knowlegde of personal data collection and usage			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Users with low knowledge of personal data collection and usage	Level of knowledge of personal data collection and usage	Correlation Coefficient	-,192**	-,167**	-,085	-,215**	-,165**
			Sig. (2-tailed)	,001	,005	,155	,000	,006
			N	279	279	279	279	279
	Users with high knowledge of personal data collection and usage	Level of knowledge of personal data collection and usage	Correlation Coefficient	,368**	,291**	,361**	,252**	,269**
			Sig. (2-tailed)	,000	,000	,000	,000	,000
			N	519	519	519	519	519

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 4.15: Here you can see the results of the bivariate correlation analysis between the degree of Knowledge Of Personal Data Collection And Usage and the Concern For Information Privacy construct and its dimensions. But now, the results were split into two groups for the analysis, one below 3.86 on the knowledge scale and one above. Spearman's correlation coefficient was used.

The results of Figure 4.15 spoke for itself. Below the cut-off point in the "low knowledgeability" region, there were significant negative correlations between the Knowledge Of Personal Data Collection And Usage versus CFIP and three of its four dimensions. Above the cut-off point in the "high knowledgeability" region,

larger and more significant positive correlations existed between the Knowledge Of Personal Data Collection And Usage versus CFIP and all of its dimensions. This proved the non-monotonic relationship between the degree of Knowledge Of Personal Data Collection And Usage versus the Concern For Information Privacy and the three dimensions Collection, Unauthorized Secondary Use and Improper Access. The Errors dimension did not wield a significant correlation in the "low knowledgeability" region.

The hypothesis H2 on this topic was "Individuals with very little or very much Knowledge Of Personal Data Collection And Usage will have more Concerns For Information Privacy than individuals with an average Knowledge Of Personal Data Collection And Usage." The above analysis confirmed this hypothesis.

There is a non-monotonic parabolic relation between the level of Knowledge Of Personal Data Collection And Usage on the one hand and an individual's Concerns For Information Privacy and the three dimensions Collection, Unauthorized Secondary Use and Improper Access on the other hand.

Even though this results seemed odd, it had a very intuitive interpretation. People are inclined to have a certain fear for the unknown and to distrust things they do not understand. The same goes for personal data collection and usage activities. When people have little knowledge of what happens with their data, they are mistrustful and are concerned for their information privacy. On the other side of the spectrum, knowing too much will also make people concerned for their information privacy. Companies and even governments have been known to infringe people's information privacy in the past, see for example the surveillance activities of the US government (Rackow, 2002), as disclosed by whistle-blower Edward Snowden. So if people have a very good understanding of the personal data collection and usage activities of companies, then they know they have something to be concerned about. Being halfway in the middle however tends to be less concerning for people, as they feel they know enough to trust companies with their information, which causes them to feel more comfortable (i.e. less concerned).

This relation is very similar to 1993 finding of MacKenzie between the involvement in knowledge production and the uncertainty in that knowledge. People close and intimately connected to the knowledge production were more unsure about their knowledge claims than those who indirectly rely on that knowledge. (MacKenzie, 1993) In the other extreme, people who were alienated from the knowledge felt that the uncertainty was much greater. A visualisation of this relation can be found in Figure 2.7. This relation was called the certainty trough. Apparently there also exists a privacy concerns trough.

4.2.6. WILLINGNESS TO DISCLOSE

In Figure 4.16 it can be seen that a surprisingly large amount of people were willing to share *"any additional available data about them"* with edX researchers, even though most people also stated to have significant concerns for their information privacy. This could be explained by two factors.

First of all, this was a beautiful example of the privacy paradox, which is the discrepancy between people's attitudes and behaviour in terms of privacy. When people are specifically asked about their opinion on information privacy, they usually state to have significant concerns. But when actual behaviour is measured, people have suddenly forgotten their opinions and often do not mind to disclose their personal information. Second, this question was framed in a very reliable manner. Every respondent was very familiar with the edX platform and had built up a certain trust with edX. Because of this, trusting their personal information to "edX researchers" was much easier, since it did not feel like they were giving their information to a stranger.

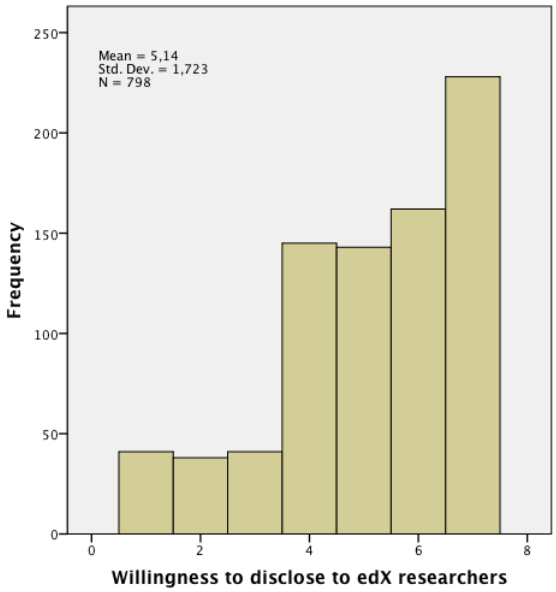


Figure 4.16: Here you can see the histogram of the responses to the statement: "I would be happy if edX researchers would use any additional available data of me to improve the quality of this research." Respondents were asked to answer on a seven point Likert scale anchored by strongly disagree (1) and strongly agree (7).

Correlations

			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Willingness to disclose to edX researchers	Correlation Coefficient	,107**	-,069	,254**	,113**	,172**
		Sig. (2-tailed)	,003	,051	,000	,001	,000
		N	798	798	798	798	798

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.17: Here you can see the correlation matrix of Willingness To Disclose vs CFIP.

In Figure 4.17 the correlation matrix between the Willingness To Disclose and Concern For Information Privacy has been posted. Here an even stranger phenomenon was detected. There appeared to be a significant positive correlation between Concern For Information Privacy and the Willingness To Disclose to edX researchers, which meant that people who were more concerned with their information privacy were more happy to share their data with edX researchers. The explanation behind this strange result may be found in the trustworthy image of edX and in the multiple dimensions of the CFIP construct.

It should be noted that the relation with the Collection dimension was, although only just insignificant with $p = 0.051$, negatively correlated with the Willingness To Disclose to edX researchers. Apparently, people who worried about Errors in their personal information and about the Unauthorized Secondary Use and Improper Access of their personal information also had more trust in edX and how they handle personal information.

Apart from this, the question was posed in an odd fashion to assess behaviour instead of intention, and the respondents were already filling out a questionnaire via edX, so apparently all respondents were willing to disclose their personal information to edX researchers. This could all have contributed to the strange result in Figure 4.17.

The previously defined hypotheses H9 was "Individuals with more Concerns For Information Privacy will be less willing to disclose their information in a MOOC setting." This hypotheses was rejected.

4.2.7. INTERNET EXPERIENCE

In Figure 4.18 you can find the correlation matrix between the Frequency Of Social Media Use, Frequency Of Online Forum Use and Average Time Spend Online Daily versus the Concern For Information Privacy construct and its dimensions.

Correlations

			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Frequency of social media use	Correlation Coefficient	-,067	-,193**	,045	-,018	,015
		Sig. (2-tailed)	,057	,000	,208	,613	,662
		N	798	798	798	798	798
	Frequency of online forum use	Correlation Coefficient	-,023	-,073*	,067	-,077*	-,019
		Sig. (2-tailed)	,516	,040	,059	,030	,583
		N	798	798	798	798	798
	Average time spend online daily in minutes	Correlation Coefficient	-,047	-,034	-,074*	-,005	,014
		Sig. (2-tailed)	,189	,339	,037	,890	,687
		N	798	798	798	798	798

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.18: Here you can see the results of the bivariate correlation analysis between the Frequency Of Social Media Use, Frequency Of Online Forum Use and Average Time Spend Online Daily versus the Concern For Information Privacy construct and its dimensions. Spearman's correlation coefficient was used.

None of the constructs appeared to have a significant correlation with the overall CFIP construct, but several significant relations did arise. The negative correlation between Frequency Of Social Media Use and the concern for Collection of personal data was very significant ($p = 0.000$), which made sense because people who often use social media, willingly share their personal information. The same goes for people who often use online forums. The developed hypotheses on this subject can be found below.

- H11A: Individuals who spend more time online will have less Concerns For Information Privacy.
- H11B: Individuals who spend more time using social media will have less Concerns For Information Privacy.
- H11C: Individuals who spend more time using online forums will have less Concerns For Information Privacy.

Because all three constructs did not yield a significant relation with the CFIP construct, all three above presented hypotheses were rejected.

The literature review of Li (2011) predicted a non-monotonic relation between internet knowledge and the Concern For Information Privacy. The Average Time Spend Online Daily was expected to be a good predictor of internet knowledge, but a scatterplot (not shown here) revealed that this non-monotonic relation between the Average Time Spend Online Daily and the CFIP was not present.

4.2.8. REALISTIC ALTERNATIVES

In Figure 4.19 you can find the correlation matrix of the Frequency Of Use Of Services Which Require The Collection Of Personal Information and the Importance Of Realistic Alternatives for privacy infringing applications versus the Concern For Information Privacy construct and its four dimensions.

The results of Figure 4.19 had an astonishing significance. Both the Frequency Of Use Of Services Which Require The Collection Of Personal Information and the perceived importance of realistic privacy respecting alternatives had a significance $p = 0.000$ with the overall Concern For Information Privacy as well as all of

Correlations

			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Frequency of use of online services which require personal data collection	Correlation Coefficient	,227**	,179**	,228**	,173**	,215**
		Sig. (2-tailed)	,000	,000	,000	,000	,000
		N	798	798	798	798	798
	Percieved importance of realistic privacy respecting alternatives	Correlation Coefficient	,484**	,523**	,247**	,309**	,313**
		Sig. (2-tailed)	,000	,000	,000	,000	,000
		N	798	798	798	798	798

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 4.19: Here you can see the results of the bivariate correlation analysis between the Frequency Of Use Of Services Which Require The Collection Of Personal Information and the Importance Of Realistic Alternatives for privacy infringing applications versus the Concern For Information Privacy construct and its dimensions. Spearman's correlation coefficient was used.

the four dimensions of concerns: Collection, Errors, Unauthorized Secondary Use and Improper Access. This could be easily understood since people who use online services which require the collection of personal data more often, will be more agitated with this collection and develop more Concerns For Information Privacy. The hypothesis H11D on this relation was "Individuals who use services which require the collection of personal data more often, will have more Concerns For Information Privacy." The above results confirmed this hypothesis.

Also, people who have high Concerns For Information Privacy will more quickly see the need for realistic privacy respecting alternatives. The hypothesis on this relation was H3, "Individuals with higher levels of Concern For Information Privacy will deem realistic alternatives to services which require the collection of personal information more important." The results from Figure 4.19 confirmed this hypothesis.

Correlations

Frequency of use of services which require the collection of personal data				Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Non-frequent users	Spearman's rho	Percieved importance of realistic privacy respecting alternatives	Correlation Coefficient	,462**	,507**	,337**	,238*	,388**
			Sig. (2-tailed)	,000	,000	,001	,021	,000
			N	93	93	93	93	93
Frequent users	Spearman's rho	Percieved importance of realistic privacy respecting alternatives	Correlation Coefficient	,490**	,519**	,219**	,297**	,269**
			Sig. (2-tailed)	,000	,000	,000	,000	,000
			N	561	561	561	561	561
Average users	Spearman's rho	Percieved importance of realistic privacy respecting alternatives	Correlation Coefficient	,429**	,447**	,222**	,386**	,375**
			Sig. (2-tailed)	,000	,000	,007	,000	,000
			N	144	144	144	144	144

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.20: Here you can see the results of the bivariate correlation analysis between the Importance Of Realistic Alternatives versus the Concern For Information Privacy construct and its dimensions, for different users which use services which require the collection of personal data with different frequencies. Spearman's correlation coefficient was used.

Hypothesis H4 was "The relation of hypothesis 3 will be affected by the Frequency Of Use Of Services Which Require The Collection Of Personal Information." To test this, the correlation relation between the Importance Of Realistic Alternatives for privacy infringing applications and the Concern For Information Privacy was retested for different sample groups of infrequent, average and frequent users of services which

require the collection of personal data. The results of this additional test can be found in Figure 4.20. These correlations were again all very significant, which allowed for a thorough analysis of the relation between the Importance Of Realistic Alternatives and the Frequency Of Use Of Services Which Require The Collection Of Personal Information. This relation was further investigated by reviewing the scatterplots for different use frequencies.

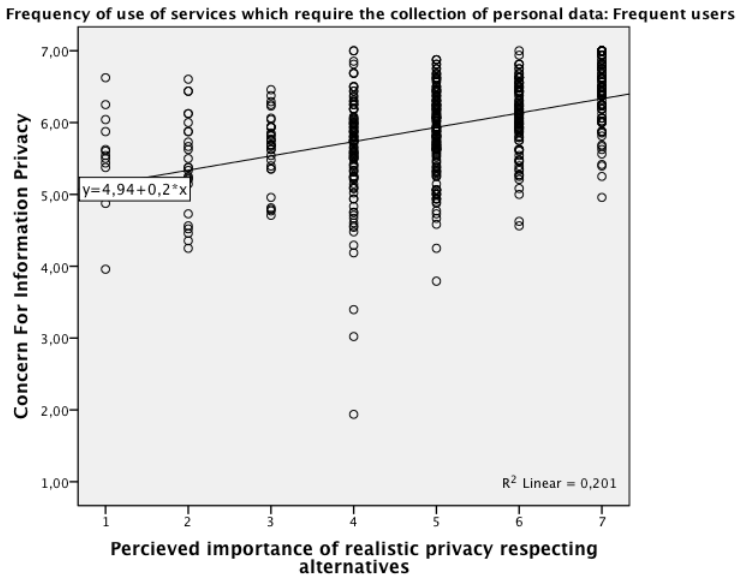


Figure 4.21: Here you can see a scatterplot of the Importance Of Realistic Alternatives versus the Concern For Information Privacy. Only the cases with a high Frequency Of Use Of Services Which Require The Collection Of Personal Information were included.

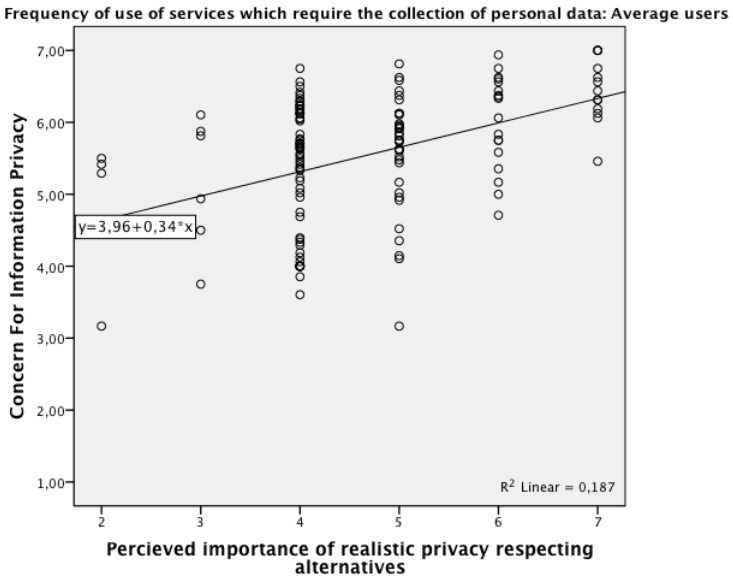


Figure 4.22: Here you can see a scatterplot of the Importance Of Realistic Alternatives versus the Concern For Information Privacy. Only the cases with an average Frequency Of Use Of Services Which Require The Collection Of Personal Information were included.

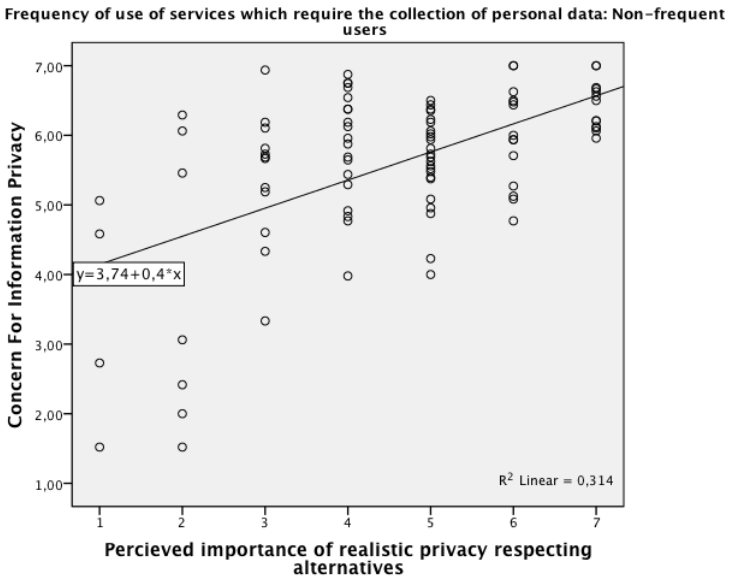


Figure 4.23: Here you can see a scatterplot of the Importance Of Realistic Alternatives versus the Concern For Information Privacy. Only the cases with a low Frequency Of Use Of Services Which Require The Collection Of Personal Information were included.

In Figures 4.21 to 4.23, you can see the scatterplots of the Importance Of Realistic Alternatives versus the CFIP for different sample groups of individuals who infrequently, averagely and frequently use services which require the collection of personal data. The scatterplots were outlined with linear approximations and the exact formula's of these lines were included in the figures. The slopes of the linear approximations were:

$$m_{infrequent} = 0.40 \tag{4.1}$$

$$m_{average} = 0.34 \tag{4.2}$$

$$m_{frequent} = 0.20 \tag{4.3}$$

These numbers and scatterplots showed that an increase in the Frequency Of Use Of Services Which Require The Collection Of Personal Information caused a decrease in the slope of relation between the Importance Of Realistic Alternatives and the CFIP. This was hypothesised in H4, "The relation of hypothesis 3 will be affected by the Frequency Of Use Of Services Which Require The Collection Of Personal Information." Therefore this hypothesis was approved.

4.2.9. INFLUENCE OF GENDER, AGE AND HIGHEST LEVEL OF EDUCATION ON CFIP

First of all, Gender did not have a significant influence on the CFIP or any of the four dimensions. The literature review of Li (2011) however listed seven researches which found that women were more concerned for information privacy than men, versus two researches which did not yield significant results.

Considering the large sample size of this study, this was an unexpected result. Apparently there were no differences in the Concerns For Information Privacy between men and women in the sample used. Perhaps this was due to the average higher education and/or the high percentage of men in the sample. The relation between Gender and CFIP was not hypothesized, so no hypothesis had to be rejected.

The relations between the Highest Level Of Education, Position In Society and Age can be found in Figure 4.24. Here can be seen that the in literature previously found relation between Age and the Concern For Information Privacy was confirmed by this research. Therefore, the hypothesis H5D "Individuals with a higher Age are more concerned for their information privacy", was confirmed.

Besides this, small but significant correlations were found between ones Position In Society and ones

concerns for Unauthorized Secondary Use and Improper Access of personal information. A separate analysis of the relation between ones Position In Society and ones Concerns For Information Privacy and the four dimensions (not shown here) was also executed on a sub-sample of financially independent individuals, but this yielded no significant result. The hypothesis H5B developed for this relation was "Individuals with a better Position In Society are more concerned for their information privacy." This hypothesis was rejected because there was no significant correlation between an individual's Position In Society and the overall Concern For Information Privacy construct.

Correlations

			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Highest level of education	Correlation Coefficient	,080*	,144**	-,005	,065	,052
		Sig. (2-tailed)	,025	,000	,879	,068	,148
		N	782	782	782	782	782
	Position in society	Correlation Coefficient	,034	-,034	,047	,072*	,077*
		Sig. (2-tailed)	,343	,336	,188	,042	,031
		N	798	798	798	798	798
	Age	Correlation Coefficient	,217**	,307**	,059	,141**	,117**
		Sig. (2-tailed)	,000	,000	,096	,000	,001
		N	798	798	798	798	798

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.24: Here you can see the results of the bivariate correlation analysis of Age, the Highest Level Of Education and the Position In Society versus the Concern For Information Privacy construct and its dimensions. Spearman's correlation coefficient was used.

And thirdly, I can deduce from Figure 4.24 that a small but significant positive correlation existed between a person's Highest Level Of Education and his/her Concerns For Information Privacy. This relation resembled the relation found by Milne *et al.* (1996) between the attitude towards privacy and education. But a relation between education and the CFIP construct was never found before. A confirmation of this result can be found in Figure 4.25 in which a K independent samples t test was executed. The hypothesis H5A developed for this relation was "Individuals with a higher level of education are more concerned for their information privacy." This hypothesis was confirmed by the presented results.

Test Statistics[a,b]

	Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Chi-Square	21,931	26,991	23,079	9,868	10,422
df	6	6	6	6	6
Asymp. Sig.	,001	,000	,001	,130	,108

a. Kruskal Wallis Test

b. Grouping Variable: Highest level of education

Figure 4.25: Here you can see the results of the k independent samples t test of the CFIP construct and its dimensions for different Highest Levels Of Education.

In Figure 4.26 you can find a bar chart of the mean Concern For Information Privacy categorized by the Highest Level Of Education.

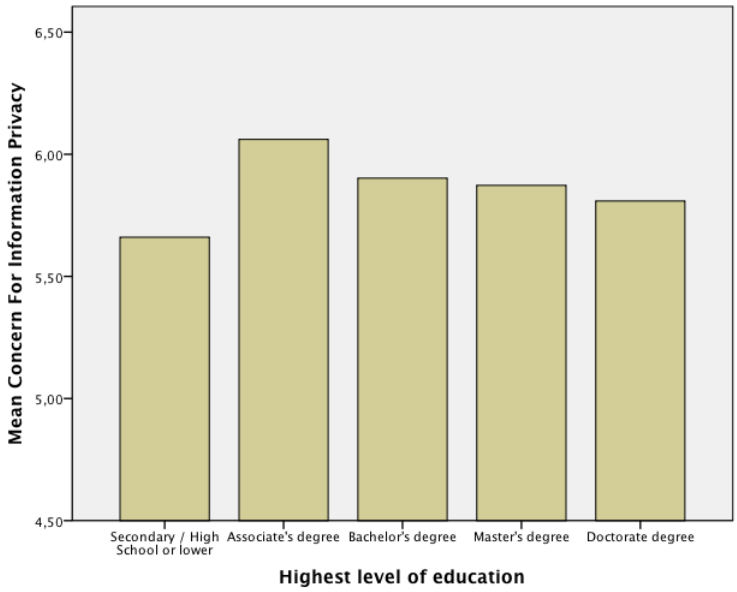


Figure 4.26: Here you can see a bar chart of the Highest Level Of Education indicating the differences in Concerns For Information Privacy.

4.2.10. SCHWARTZ' HUMAN VALUES

In Figure 4.27, you can find the results of the bivariate correlation analysis between the five selected universal human values as defined by Schwartz (1994) and the CFIP construct and its four dimensions.

Correlations

			Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Spearman's rho	Universalism	Correlation Coefficient	,201**	,101**	,209**	,182**	,202**
		Sig. (2-tailed)	,000	,004	,000	,000	,000
		N	798	798	798	798	798
	Self-direction	Correlation Coefficient	,264**	,106**	,265**	,222**	,264**
		Sig. (2-tailed)	,000	,003	,000	,000	,000
		N	798	798	798	798	798
	Stimulation	Correlation Coefficient	,171**	,020	,216**	,161**	,189**
		Sig. (2-tailed)	,000	,577	,000	,000	,000
		N	798	798	798	798	798
	Hedonism	Correlation Coefficient	,061	,051	,065	,067	,043
		Sig. (2-tailed)	,084	,147	,068	,057	,221
		N	798	798	798	798	798
	Security	Correlation Coefficient	,150**	,014	,222**	,082*	,190**
		Sig. (2-tailed)	,000	,686	,000	,020	,000
		N	798	798	798	798	798

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.27: Here you can see the results of the bivariate correlation analysis of five of the ten human values defined by Schwartz (Universalism, Self-Direction, Stimulation, Hedonism and Security) versus the Concern For Information Privacy construct and its dimensions. Spearman's correlation coefficient was used.

In the above figure can be seen that apart from Hedonism, the other four tested universal human values all had a significant positive correlation with the Concern For Information Privacy. Apparently, people with

higher levels of Universalism, Self-Direction, Stimulation and Security had higher Concerns For Information Privacy. The relation with the human value Security was to be expected, but the other three results are interesting.

Apart from the correlations with the overall CFIP construct, Universalism and Self-Direction had positive correlations with all four of the separate dimensions of CFIP and Stimulation and Security had positive correlations with three of the four separate dimensions of CFIP.

Because this part of the research was expected to have insufficient reliability, no hypotheses were developed. So no hypotheses were accepted or rejected on the relations between human values and the CFIP. Despite this, the above findings gave a valuable initial exploration of the relations between human values and the CFIP and its dimensions.

4.2.11. ADDITIONAL FINDINGS

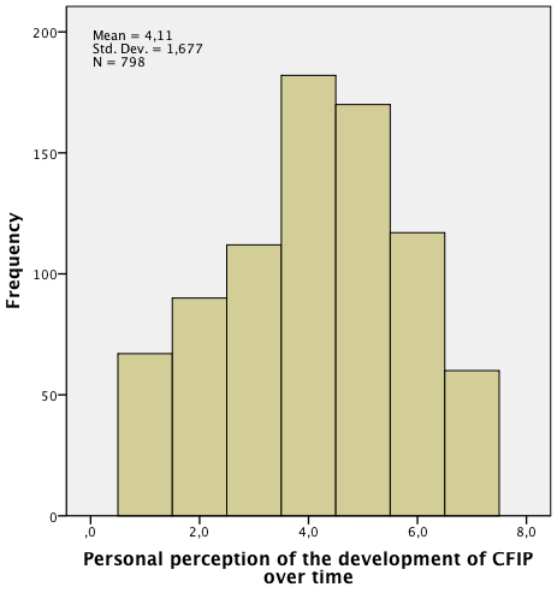


Figure 4.28: Here you can see the histogram of the single question item about an individual's own perception of it's personal development of it's view on information privacy. What stood out here was the large standard deviation of $\sigma = 1.677$.

The statement that accompanied the histogram in Figure 4.28 was "Over the past ten years my attitude towards information privacy has become more and more tolerant." The mean of this questions was $\mu = 4.11$, so society itself tended to think that its view on information privacy stayed about the same or very slightly increased during the past decade (since this was only 0.0656 standard deviations from the neutral center point):

$$\mu - \mu_{neutral} = \frac{\mu - \mu_{neutral}}{\sigma} \cdot \sigma = \frac{4.11 - 4}{1.677} \cdot \sigma = 0.0656 \cdot \sigma$$

This was in accordance with the result in study 1, in which a similar slight but non-significant positive correlation existed between the CFIP construct and time over the past twenty years.

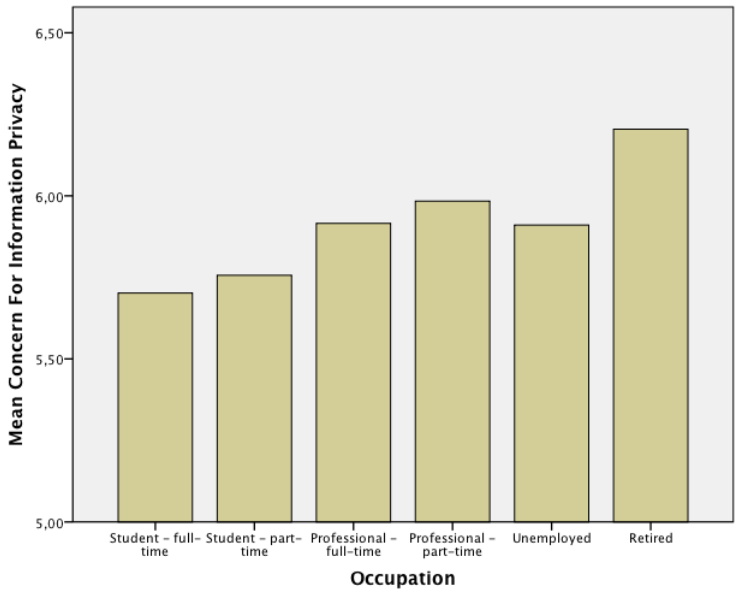


Figure 4.29: Here you can see a bar chart of people with different a Occupation and the corresponding means of Concerns For Information Privacy.

In Figure 4.29 you can find a bar chart of the mean Concern For Information Privacy which compares people with different Occupations. Several Occupations were left out of consideration because of a low amount of respondents, which could give an unreliable and skewed image. Occupations with less than 15 respondents were left out of consideration here. What immediately stood out was the discrepancy between students and other Occupations. This discrepancy was analysed by executing an independent samples t test between full time students and full time professionals. The descriptives of this analysis can be found in Figure 4.30 and the results of the analysis can be found in Figure 4.31.

	Occupation	N	Mean	Std. Deviation	Std. Error Mean
Concern For Information Privacy	Student – full-time	339	5,7019	,84178	,04572
	Professional – full-time	257	5,9157	,79864	,04982
Collection dimension	Student – full-time	339	4,8407	1,19187	,06473
	Professional – full-time	257	5,3959	1,13132	,07057
Errors dimension	Student – full-time	339	5,4993	1,00654	,05467
	Professional – full-time	257	5,4728	1,14019	,07112
Unauthorized secondary use dimension	Student – full-time	339	6,2367	1,00948	,05483
	Professional – full-time	257	6,4465	,88101	,05496
Improper access dimension	Student – full-time	339	6,2311	1,01502	,05513
	Professional – full-time	257	6,3476	,92949	,05798

Figure 4.30: Here you can see the means and standard deviations for the groups of full-time students and full-time professionals for the CFIP construct and its four dimensions.

Independent Samples Test

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Concern For Information Privacy	Equal variances assumed	-3,138	594	,002	-,21375	,06811	-,34752	-,07999
	Equal variances not assumed	-3,161	565,172	,002	-,21375	,06762	-,34656	-,08094
Collection dimension	Equal variances assumed	-5,756	594	,000	-,55521	,09645	-,74464	-,36578
	Equal variances not assumed	-5,798	565,056	,000	-,55521	,09576	-,74330	-,36711
Errors dimension	Equal variances assumed	,301	594	,764	,02650	,08818	-,14669	,19969
	Equal variances not assumed	,295	512,388	,768	,02650	,08971	-,14974	,20274
Unauthorized secondary use dimension	Equal variances assumed	-2,652	594	,008	-,20977	,07909	-,36510	-,05444
	Equal variances not assumed	-2,702	582,301	,007	-,20977	,07763	-,36224	-,05731
Improper access dimension	Equal variances assumed	-1,439	594	,151	-,11653	,08098	-,27557	,04251
	Equal variances not assumed	-1,457	573,251	,146	-,11653	,08001	-,27367	,04061

Figure 4.31: Here you can see the results of the independent t test between the groups of full-time students and full-time professionals for the CFIP construct and its four dimensions.

From the above analysis was concluded that full time students were apparently less concerned than full time professionals about the Collection of their personal data, the Unauthorized Secondary Use of their personal data and about their information privacy in general.

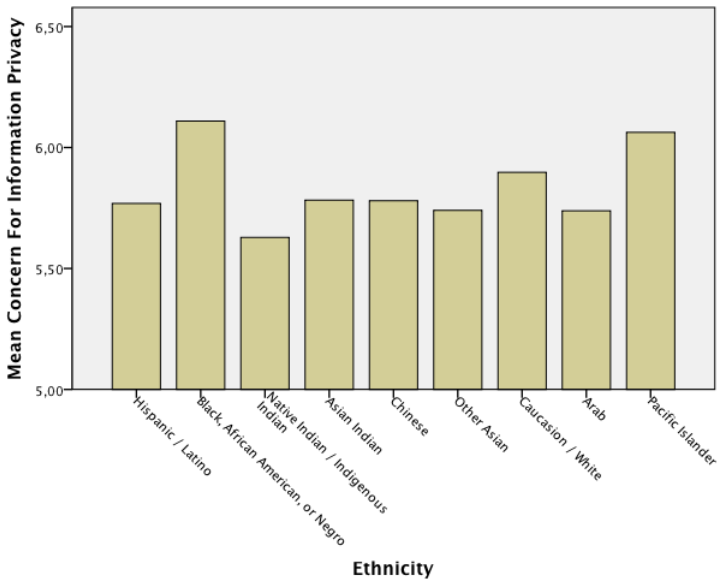


Figure 4.32: Here you can see a bar chart of people with different ethnicities and the corresponding means of Concerns For Information Privacy.

Test Statistics[a,b]

	Concern For Information Privacy	Collection dimension	Errors dimension	Unauthorized secondary use dimension	Improper access dimension
Chi-Square	7,821	42,354	28,220	9,642	11,937
df	13	13	13	13	13
Asymp. Sig.	,855	,000	,008	,723	,533

a. Kruskal Wallis Test

b. Grouping Variable: Ethnicity

Figure 4.33: Here you can see the K Independent samples test CFIP grouped by ethnicity.

In Figure 4.32 you can see a bar chart of the Concern For Information Privacy for people of different ethnicities. Several ethnicities were left out of consideration because of a low amount of respondents, which could give an unreliable and skewed image. Ethnicities which had less than 15 respondents were left out of consideration here. To find out whether differences in ethnicity corresponded to differences in the Concern For Information Privacy, a K independent samples test was executed, this can be seen in Figure 4.33. From this analysis was concluded that there were significant differences in the concern for the Collection of and Errors in personal information between ethnicities, but not in the general CFIP construct. This was in accordance with the geographical cluster analysis which was reported on page 50.

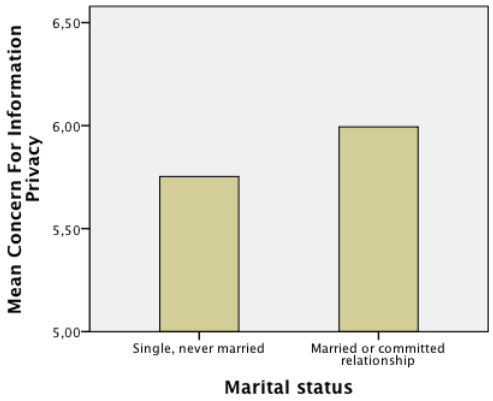


Figure 4.34: Here you can see a bar chart which emphasized the difference in Concerns For Information Privacy between people with and without a relationship.

Independent Samples Test

		t-test for Equality of Means				
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Concern For Information Privacy	Equal variances assumed	-3,746	774	,000	-,24045	,06419
	Equal variances not assumed	-3,795	383,556	,000	-,24045	,06336

Figure 4.35: Here you can see the results of the independent samples t test of the Concern For Information Privacy construct between singles and people who were married or were in a committed relationship.

The final relation which was investigated was the relation between Marital Status and the CFIP. The bar chart can be found in Figure 4.34, and the independent samples t test can be found in Figure 4.35. It turned out that people who were married were significantly more concerned for their information privacy than single individuals. This was expected and was partly be explained by the fact that the average married person had a higher Age than the average single person ($\mu_{Age,married} = 39.5$ vs $\mu_{Age,single} = 23.6$).

4.2.12. SIMPLE CONCEPTUAL MODELS

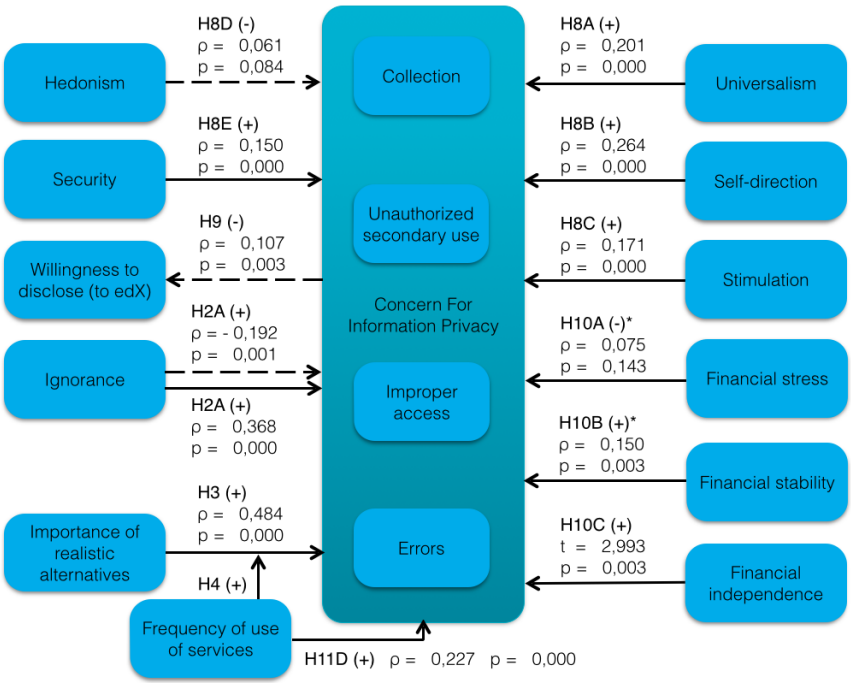


Figure 4.36: Here you can see the first conceptual model of study 2 which was developed in the research framework chapter, but now with the magnitudes and significances of the results from the correlation analyses integrated and also with indications which hypotheses were supported and which were not. Note that the relations indicated with an asterisk were tested on a sub-sample of financially independent individuals.

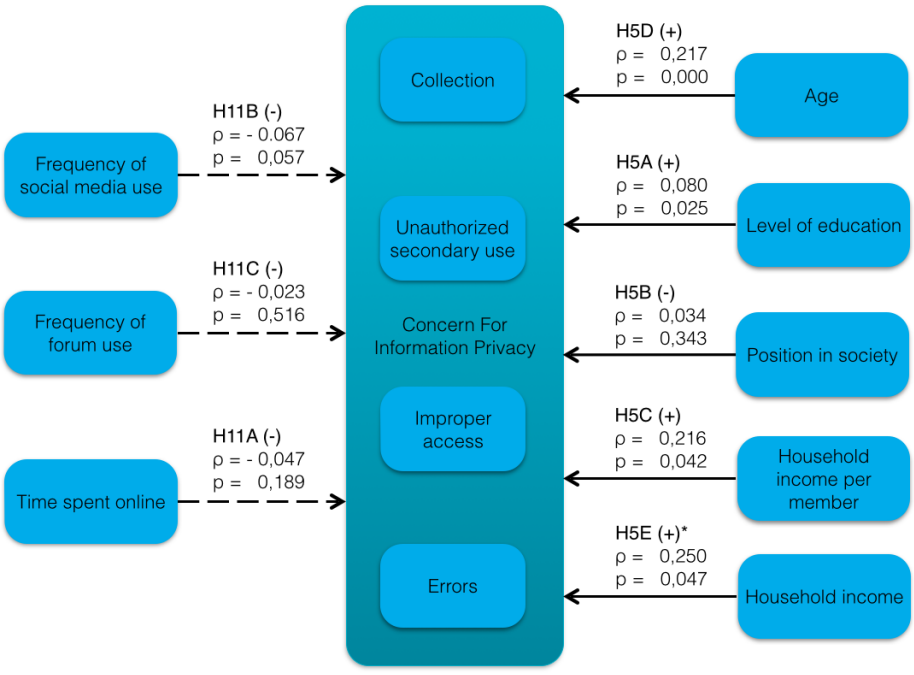


Figure 4.37: Here you can see the second conceptual model of study 2 which was developed in the research framework chapter, but now with the magnitudes and significances of the results from the correlation analyses integrated and also with indications which hypotheses were supported and which were not. Note that the relation indicated with an asterisk was tested on a sub-sample of financially independent individuals.

In Figures 4.36 and 4.37 you can find the results of study 2 summarized in the previously defined conceptual models.

4.2.13. MEDIATING AND MODERATING CONCEPTUAL MODEL

The models of the previous subsection only consisted of direct correlations, but it was also statistically possible to test for mediation and/or moderation. A thorough analysis of the data found a more complex model which suited the data accurately. This model consisted of several mediating and moderating relations and was found by using the advanced process macro software which was developed by Andrew Hayes (Hayes, 2012). This open source software allowed for an easy trail and error testing of 74 different models. This resulted in the model based on Hayes' "Model 17", which you can find in Figure 4.38. You can find the original conceptual and statistical model as created by Hayes in Figure 4.39. The output text which resulted from computing the analysis of the standardized and centered model 17 can be found in appendix E.

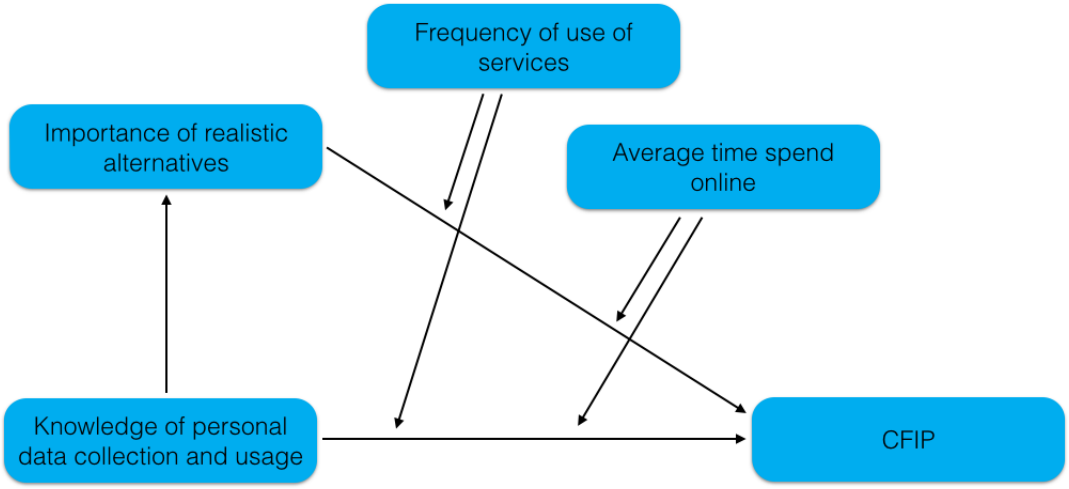


Figure 4.38: Here you can see the complex conceptual model of study 2 which was developed with the process marco software of Andrew Hayes.

In Figure 4.38 you can see how the relation between the Knowledge Of Personal Data Collection And Usage and the Concern For Information Privacy was elaborated on. In the model the Importance Of Realistic Alternatives served as a mediating variable. Both relations with the Concerns For Information Privacy were than moderated by the Frequency Of Use Of Services Which Require The Collection Of Personal Information and the Average Time Spend Online Daily. This conceptual model gave a more profound insight into how the constructs "Knowledge Of Personal Data Collection And Usage", "Importance Of Realistic Alternatives", "Frequency Of Use Of Services Which Require The Collection Of Personal Information" and "Average Time Spend Online Daily" all directly or indirectly affected an individuals' Concern For Information Privacy. This was considered to give a better representation of the interactions of these constructs than the more superficial bivariate correlation analyses of the previous sections. How this mechanism presented in Figure 4.38

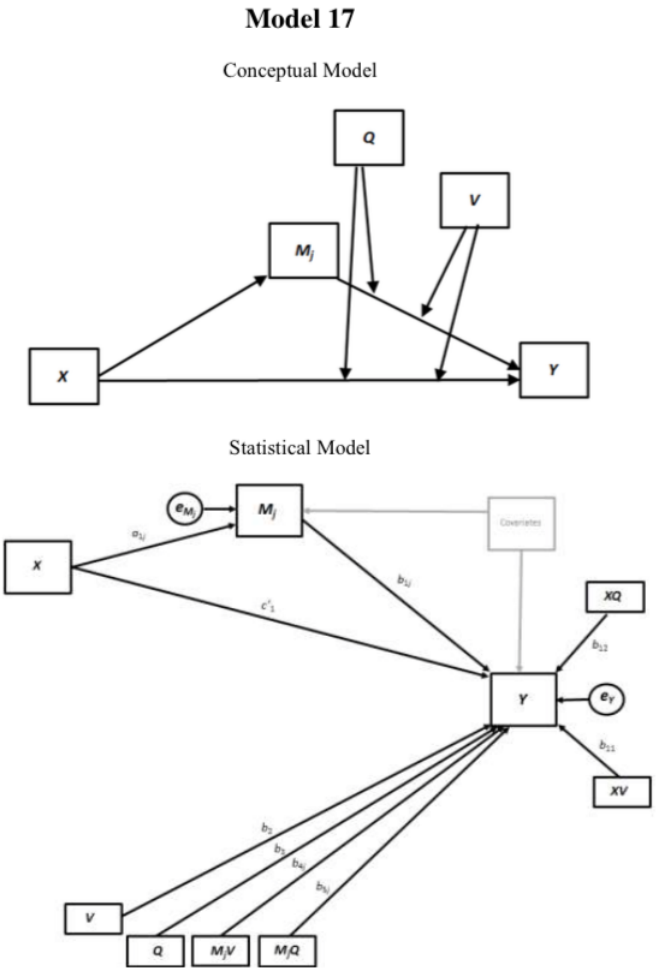


Figure 4.39: Here you can see the original conceptual and statistical model 17 of Andrew Hayes.

explained the non-monotonic parabolic relation between "Knowledge Of Personal Data Collection And Usage" and "Concern For Information Privacy" was not directly clear.

In the first instance the Highest Level Of Education was added as a covariate, but turned out to be insignificant. In the previous analysis, the respondents who indicated to have an "other" Highest Level Of Education were treated as missing values and were left out of the analysis list-wise. Therefore, to prevent a drop in sample size, the Highest Level Of Education variable was left from the previous analysis completely.

Another possibility which seemed intuitive was for the constructs "Frequency Of Use Of Services Which Require The Collection Of Personal Information" and "Average Time Spend Online Daily" to also have a moderating effect on the relation between "Knowledge Of Personal Data Collection And Usage" and "Importance Of Realistic Alternatives". This exact model was not available in Hayes' process macro, but another model (model 67) allowed for the "Frequency Of Use Of Services Which Require The Collection Of Personal Information" factor to also moderate the previously specified relation. This however resulted in an unstable model in which the mediating relation of "Importance Of Realistic Alternatives" was not significant, so this possibility was omitted.

4.3. SHORT SUMMARY OF RESULTS

The objective of this study was to investigate the influence of several new and old factors which influence the Concern For Information Privacy construct and its dimensions. This was done by executing a quantitative survey based research. This study concluded that MacKenzie's certainty trough applies to the Knowledge Of Personal Data Collection And Usage, and the Concern For Information Privacy. A significant conceptual model was developed to further explain this relation with respect to other determinants. And several relations between demographics and the CFIP were found, for example between the Highest Level Of Education, Financial Stability and Household Income Per Household Member and the CFIP construct. Several human values also seemed to be of significance with respect to the CFIP.

5

STUDY 3

5.1. METHODOLOGY

The goal of study 3 was to investigate whether the spheres of justice theory applies to exchanges of personal information. This was done quantitatively by asking people their opinion on several inter- and intra-sphere exchanges of personal information. This study also explored the existence of sub-spheres and the effect of the consequence of a personal data exchange on its perceived appropriateness.

5.1.1. RESEARCH OBJECTIVE

One of the take-aways from study 2 was the increase of people's concerns for the Improper Access of their personal information. A critical question that arose from this result was: *when is the access to personal information considered improper?* The spheres of informational justice theory (Van den Hoven, 1999) provides an explanation to why and when exchanges of personal data are considered unjust, but this theory has never been tested quantitatively before. The goal of this study was to give quantitative insight into this theory and to elaborate on the notions of sub-spheres and the influence of consequence. Therefore the main objectives of this third study were:

- Quantitatively prove the spheres of informational justice theory
- Quantitatively prove that the consequence of a personal data exchange affects how appropriate it is considered.
- Quantitatively prove the existence of sub-spheres in the context of personal information exchanges

5.1.2. SAMPLE

Similar to study 2, I would like the population to be the entire global society. This would allow for the research to be truly generalizable. Fortunately another way was found to obtain this global sample via edX. By selecting a sub-sample of the sample in study 2, I reached the same type of sample for study 3. How this acquisition was executed can be found in the following subsection.

5.1.3. DATA ACQUISITION

The questionnaire of study 2 included the following question:

We are doing a lot of exciting research at Delft University of Technology. Do you want to be a pioneer and participate in one of our research projects? Let us know by clicking yes below, and we may send an email to you after the course. As a bonus, we will also include some fresh research results.

The participants which reacted positively to this question were sent an email with an invitation to participate in the research for study 3.

5.1.4. FILTERING

In the initial emailing 532 emails were sent, of which 1 bounced, so 531 participants were invited to fill out the survey. After three weeks and two reminders, 137 respondents started the survey and 108 respondents fully completed the survey. This results in a response rate of 20.3 %. All of the 108 respondents fully completed the core section of the questionnaire, but several left some questions blank in the edX specific section. These respondents were included into the analysis and missing data was handled list-wise. This was not a problem because the maximum amount of missing values per questions was only 2.

5.1.5. QUESTIONNAIRE

The complete questionnaire which was used for study 3 can be found in appendix G. Below you can find all the topics and constructs which were measured by the questionnaire. For each topic it was outlined why it is integrated into the questionnaire and the specific constructs which were measured were explained. All but one constructs were measured by asking the respondents how appropriate they consider a certain exchange of personal data. The respondents were asked to give their answers a seven point Likert scale anchored by "Very inappropriate" and "Very appropriate".

APPROPRIATENESS OF INTRA-SPHERE PERSONAL DATA EXCHANGES

The Appropriateness Of Intra-Sphere Personal Data Exchanges was measured for the five included informational spheres: medical, governmental, educational, commercial and financial. This allowed to see the differences in the mean perceived Appropriateness Of Intra-Sphere Personal Data Exchanges between different spheres and served as a benchmark for the Appropriateness Of Intra-Sphere Personal Data Exchanges in general. Please consult appendix G for the exact list of questions.

APPROPRIATENESS OF INTER-SPHERE PERSONAL DATA EXCHANGES

The Appropriateness Of Inter-Sphere Personal Data Exchanges (i.e. exchanges between different spheres) was measured by asking respondents how appropriate they considered all possible combinations of inter-sphere personal data exchanges. This total of $5 \cdot 4 = 20$ questions gave an indication of the appropriateness of all different combinations as well as set the benchmark for the Appropriateness Of Inter-Sphere Personal Data Exchanges. Please consult appendix G for the exact list of questions.

APPROPRIATENESS OF PERSONAL DATA EXCHANGES WITH DIFFERENT CONSEQUENCES

The appropriateness of personal data exchanges with different consequences was measured in six different ways and the questions were formed in an anecdotal manner. Because of this, the questions were only formed in an educational and medical context, since these contexts allowed for realistic and empathetic anecdotes. The constructs measured were:

- Appropriateness Of Personal Data Exchanges With A Personal Negative Consequence
- Appropriateness Of Personal Data Exchanges With A Personal Positive Consequence
- Appropriateness Of Personal Data Exchanges With A Societal Positive Consequence

What was meant with these constructs was the consequence the exchange of personal data had on the individual or on society at large and whether this consequence was positive or negative. The constructs above were measured for both inter-sphere and intra-sphere exchanges, hence the six different ways. Please consult appendix G for the exact list of questions.

The case of the perceived appropriateness of an exchange of personal data with a negative consequence on society was excluded from this research simply because this is never promoted this way. If an exchange of personal data with a consequence on society is presented, potential negative consequences are never pointed out. So if an anecdotal question would be set up in this case, the question would seem odd and unrealistic.

These constructs served as indications for the average Appropriateness Of Personal Data Exchanges With A Personal Negative Consequence, A Personal Positive Consequence and A Societal Positive Consequence. So please note that these constructs were only indicators based on anecdotal examples in the medical and educational sphere. The constructs were operationalised quickly and easily to create an interesting estimation of how the consequences influenced the perceived appropriateness. There was no elaborate testing to confirm the validity, generalizability and reliability of these constructs.

APPROPRIATENESS OF INTRA-SUB-SPHERE PERSONAL DATA EXCHANGES

The goal of measuring the Appropriateness Of Intra-Sub-Sphere Personal Data Exchanges was to prove the existence of sub-spheres in which the appropriateness of an exchange is considered even higher (more appropriate) than inside the "normal" sphere. For this purpose, the appropriateness of a specific personal data exchange in three different contexts was measured in the intra-sub-sphere, intra-sphere and inter-sphere cases in an anecdotal fashion. For this part of the research the following constructs were measured separately by using anecdotal questions:

- Appropriateness Of Intra-Sub-Sphere Personal Data Exchanges
- Appropriateness Of Intra-Sphere Personal Data Exchanges
- Appropriateness Of Inter-Sphere Personal Data Exchanges

Only three contexts were used, instead of all five which were considered in this research, to constrain the length of the questionnaire. The used contexts were the medical, educational and financial context.

Apart from proving the existence of sub-spheres, this topic also allowed for a second verification of hypothesis H13 "The average appropriateness of personal data exchanges between spheres is lower (less appropriate) than the average appropriateness of personal data exchanges in the same sphere," but this time by using anecdotal questions. Please consult appendix G for the exact list of questions.

PERCEPTION OF PEOPLE ON THE PRIVACY CALCULUS

The privacy calculus is a theory in which consumers acknowledge the value of their personal information and sometimes exchange this in return for services. Although widely accepted, there is no empirical evidence yet that consumers actually perceive this exchange as it is described in the privacy calculus. The questions asked on this topic in the questionnaire aimed to map out the actual perceptions of consumers and verify the privacy calculus. Because the privacy calculus is already so accepted and because of its high plausibility, this verification was not hypothesized. Please consult appendix G for the exact list of questions.

EDX RELATED CONSTRUCTS

The below four constructs were all related to edX. Resulting from the collaboration with DelftX, some research was also done into the appropriateness of edX related uses of its students' personal data. More detailed, research was done into:

- the Appropriateness Of Different Uses Of Personal Data Of EdX;

- the Influence Of Intention;
- the Influence Of Data Type;
- the Influence Of Data Recipient.

First, to get a general impression of the Appropriateness Of Different Uses Of Personal Data Of EdX, 19 different scenarios were presented to the respondents. To get a more thorough insight into the opinions of edX users, the influence of three different factors was explored: Intention, Data Type and Data Recipient. To measure the influence of these factors, scenarios with different Intentions, Data Types and Data Recipients were presented to the respondents. Identical to the previously discussed constructs, respondents were asked to indicate the level of appropriateness for all different scenarios. Please consult appendix G for the exact list of questions.

DEMOGRAPHICS

Because the sample used in this study was a sub-sample of study 2, the respondents could be linked back to their demographics which were known from study 2. Therefore it was not necessary to include any demographics in this questionnaire.

5.2. RESULTS

5.2.1. DEMOGRAPHICS

As discussed in the methodology section above, the sample for this study was drawn from the larger sample of study 2. Therefore the demographics were very similar to the demographics of the sample of study 2. You can find them presented in Table 5.1. As before, there was a high percentage of male respondents and a small augmentation to lower age, higher education and a student occupation.

Profile	Items	Frequency	Percentage
Gender	Male	94	87.0
	Female	14	13.0
	Other	0	0.0
Age	< 20	17	15.7
	20 - 29	50	46.3
	30 - 39	18	16.7
	40 - 49	12	11.1
	50 - 59	7	6.5
	> 59	4	3.7
Education level	Doctorate	6	5.6
	Master's degree	19	17.6
	Bachelor's degree	42	38.9
	Associate's degree	4	3.7
	High school	32	29.6
	Less than high school	2	1.9
	Other	3	2.8
Ethnicity	Caucasion / White	39	36.1
	Asian Indian	27	25.0
	Hispanic / Latino	17	15.7
	Other Asian	10	9.3
	African American / Black	3	2.8
	Arab	1	0.9
	Native Indian	3	2.8
	Other	8	7.4
Occupation	Student	47	43.6
	Professional	40	37.0
	Unemployed	16	14.8
	Other	5	4.6

Table 5.1: Here you can find the demographics of the sample of study 3.

5.2.2. PUTTING THE SPHERES TO THE TEST

In Figure 5.1 you can find the paired samples t test which analysed the difference between the Appropriateness Of Intra-Sphere Personal Data Exchanges and the Appropriateness Of Inter-Sphere Personal Data Exchanges. This test was selected because it was designed to determine differences in means of two interval variables of a single sample, which was exactly what was needed for this analysis. In Figure 5.2 you can see the first conceptual model developed for this study, but now with the results of the analysis included. And in Figure 5.3 this result is visually displayed. The absolute difference of mean perceived appropriateness between inter- and intra-sphere personal data exchanges was $\Delta\mu = 0.43241$. This was fairly small considering this was not even half a point on the seven point Likert scale which was used. But on the other hand, the significance of this difference was an astounding $p = 0.000$.

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
ISAVG - BSAVG	,43241	1,00655	,09686	,24040	,62441	4,464	107	,000

Figure 5.1: Here you can find the results of the paired samples test between the intra-sphere and inter-sphere mean appropriateness of personal data exchanges. This result proved with a $p = 0.000$ significance that exchanges of personal data between different spheres were deemed less appropriate than exchanges of personal data inside the same sphere.

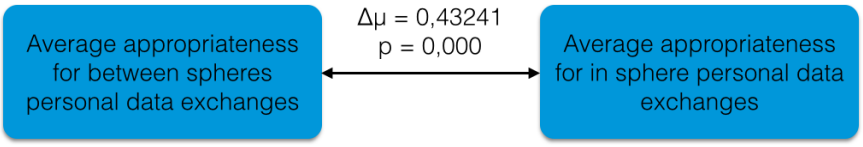


Figure 5.2: The first conceptual model developed for study 3 with the results included.

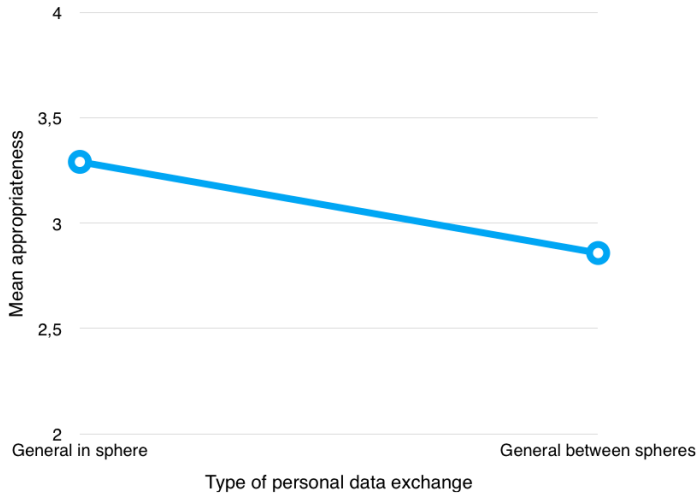


Figure 5.3: Here you can find the mean appropriateness of personal data exchanges in the same sphere and between different spheres. This proved that the spheres of justice theory is indeed applicable to exchanges of personal data.

Although this did give empirical evidence to the spheres of informational justice theory, it did not claim that intra-sphere exchanges of personal data are appropriate. The measured mean Appropriateness Of Intra-Sphere Personal Data Exchanges was $\Delta\mu = 3.2907$, which was somewhere between "3 = Somewhat inappropriate" and "4 = Neutral".

So maybe the informational spheres of justice theory should be adjusted to claim that intra-sphere data exchanges are only considered "less inappropriate" instead of the previously proposed black and white distinction between "intra-sphere = appropriate" vs "inter-sphere = inappropriate". A further analysis of the data elaborated on the superficial results found above.

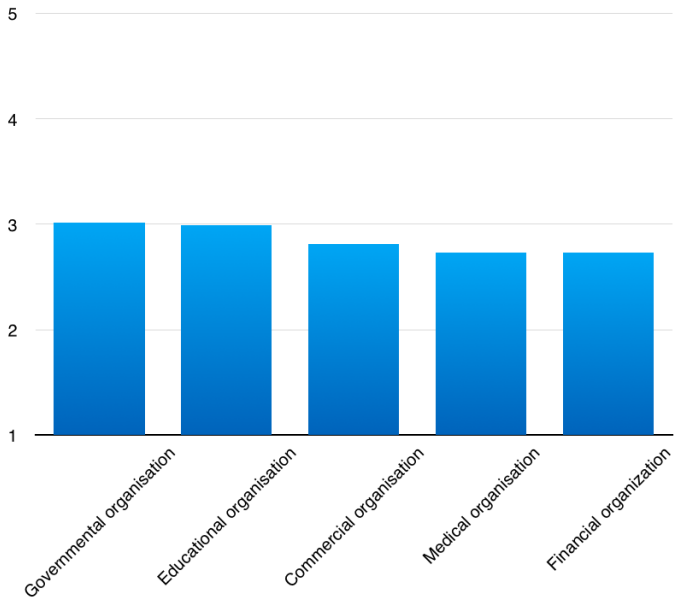


Figure 5.4: Here you can find the mean appropriateness of personal data exchanges between spheres with different types of organisations initiating the exchange. A remarkable result was that exchanges initiated by the government are considered least inappropriate.

In Figure 5.4 you can see the Appropriateness Of Inter-Sphere Personal Data Exchanges initiated by different organisation types. Intra-sphere exchanges of personal data were excluded in this analysis, so only the inter-sphere cases were included, meaning that each data point was based on the appropriateness of personal data exchanges with the other four spheres. The difference in mean appropriateness between the most and least appropriate spheres, organisational and financial respectively, was small and had an absolute value of $\Delta\mu = 0.28935$ with a significance of $p = 0.026$. Actually, the only significant differences in the average Appropriateness Of Inter-Sphere Personal Data Exchanges were between the governmental or educational sphere initiating the exchange versus the medical or financial sphere initiating the exchange. So four of the ten possible combinations of mean appropriateness were statistically different.

Despite these significant differences, the Appropriateness Of Inter-Sphere Personal Data Exchanges initiated by different organisation types were all in about the same range of appropriateness, which was around the scale point "Somewhat inappropriate".

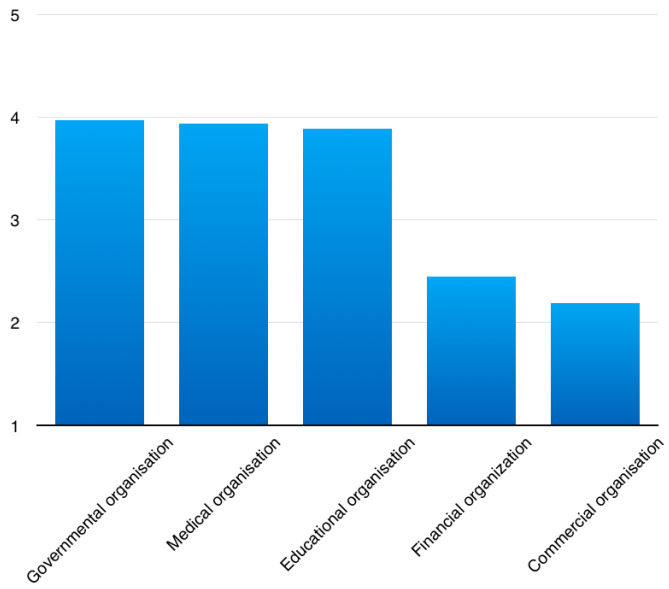


Figure 5.5: Here you can find the mean appropriateness of personal data exchanges with different types of organisations inside their own sphere. A very remarkable result was that exchanges in the financial and commercial sphere were considered inappropriate inside their own spheres.

Similar to Figure 5.4, in Figure 5.5 you can see the Appropriateness Of Intra-Sphere Personal Data Exchanges. Here I found an extraordinary large difference in the mean perceived Appropriateness Of Intra-Sphere Personal Data Exchanges for different spheres! Where the Appropriateness Of Intra-Sphere Personal Data Exchanges of the governmental, medical and educational spheres was considered around "4 = neutral", the Appropriateness Of Intra-Sphere Personal Data Exchanges of the financial and governmental sphere was considered between "3 = Somewhat inappropriate" and "2 = Inappropriate". This difference of 1,5 point on the Likert scale was huge compared to the differences seen before in this study. Very remarkable was that the intra-sphere exchanges of personal data in spheres which were linked to the public domain (governmental, medical and educational) were considered much more appropriate than those linked to the private domain (commercial and financial). This was confirmed when I looked at the perceived Appropriateness Of Inter-Sphere Personal Data Exchanges from the commercial sphere to the financial sphere and visa versa. These were $\mu = 2.44$ and $\mu = 1.96$, respectively. Which were in the same range as the perceived Appropriateness Of Intra-Sphere Personal Data Exchanges in the commercial and financial spheres.

The intra-sphere personal data exchanges in the financial and commercial spheres were considered so inappropriate that exchanges outside the sphere were on average considered even more appropriate! The difference in appropriateness between inter- and intra-sphere exchanges of personal data for the financial sphere was $\Delta\mu = 0.27546$ with a significance of $p = 0.037$. The difference in appropriateness between inter- and intra-sphere exchanges of personal data for the commercial sphere was $\Delta\mu = 0.062269$ with a significance of $p = 0.000$. This disproved the spheres of informational justice theory for the financial and commercial spheres, because these results said that the intra-sphere exchanges of personal data in these spheres are *less appropriate* than the corresponding inter-sphere exchanges, exactly the opposite of what was predicted by the spheres of informational justice theory!

> *"This disproved the spheres of informatinoal justice theory for the financial and commercial spheres."*

The hypothesis H13, which was developed to test whether the spheres of informational justice theory would hold up in reality, was "The average appropriateness of personal data exchanges between spheres is lower

(less appropriate) than the average appropriateness of personal data exchanges in the same sphere." The results stated in Figure 5.2 would approve this hypothesis. But from the analysis above was concluded that this hypothesis could only be confirmed for the medical, educational and governmental spheres.

Please note that because of the discrepancy found above, in the remainder of this thesis I will often refer to the governmental, educational and medical spheres as the "public domain" and to the financial and commercial spheres as the "private domain". I defined these two groups like this because in the governmental, educational and medical spheres most organisations are public and in the financial and commercial spheres most organisations are private. However, this is not the focus of this research and I only made these definitions for convenience.

5.2.3. INFLUENCE OF CONSEQUENCE

In Figures 5.6 and 5.7 you can find the results of the tests which were designed to map out the influence of the consequence of an exchange of personal data. For the two tests below both the inter-sphere and intra-sphere data were accumulated into the three variables. This was done because highlighting this discrepancy was not the goal of these tests, the goal was to find out whether the consequences of a personal data exchange influenced the perceived appropriateness of the exchange. By adding up the inter-sphere and intra-sphere data, the effect of the spheres was cancelled out, so the pure effect of the consequences could be analysed.

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
SOCPOSAVG - PERNEGAVG	1,13194	1,22592	,11796	,89809	1,36579	9,596	107	,000

Figure 5.6: Here you can find the results of the paired samples test between the mean appropriateness of data exchanges with societal positive and personal negative consequences.

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
PERPOSAVG - SOCPOSAVG	,41204	,76222	,07334	,26664	,55743	5,618	107	,000

Figure 5.7: Here you can find the results of the paired samples test between the mean appropriateness of data exchanges with societal positive and personal positive consequences.

In Figure 5.8 you can see the second conceptual model of this study, again with the results of the analysis included. And in Figure 5.9 this result is visually displayed. The absolute difference between the Appropriateness Of Personal Data Exchanges With A Personal Negative Consequence and the Appropriateness Of Personal Data Exchanges With A Societal Positive Consequence was $\Delta\mu = 1.13194$. This was a huge difference of more than an entire point on the seven point Likert scale. The absolute difference between the Appropriateness Of Personal Data Exchanges With A Societal Positive Consequence and the Appropriateness Of Personal Data Exchanges With A Personal Positive Consequence was $\Delta\mu = 0.41204$. This was a much smaller difference but still fairly large considering this was the difference between two positive outcomes. Both differences in appropriateness had again amazing significances of $p = 0.000$.

This meant that hypothesis H14 was also confirmed, since this hypothesis was "The average appropriateness of personal data exchanges is dependent on the consequence of the exchange, both in and between spheres. From least to most appropriate: personal negative consequence, societal positive consequence, personal positive consequence."

What could be misleading here is that this difference in perceived appropriateness seemed to be solely accounted for by the difference in outcome. This is partly true, but respondents had most likely regarded the intentions of the party initiating the personal data exchange. So these large differences between the perceived appropriateness of personal data exchanges with different outcomes also nicely reflected the Influence Of Intention.

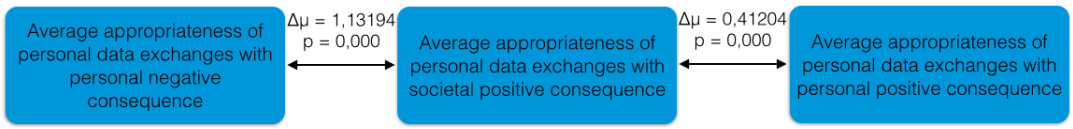


Figure 5.8: The second conceptual model developed for study 3 with the results included. The goal of this model was to prove that the consequences of a personal data exchange influence its perceived appropriateness.

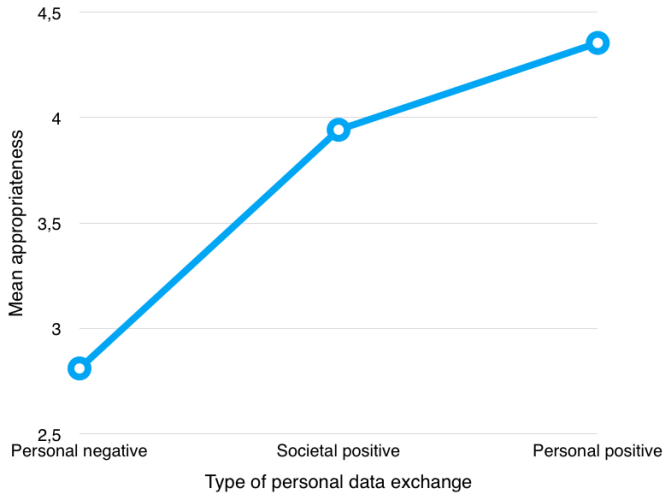


Figure 5.9: Here you can find the mean perceived appropriateness of personal data exchanges with societal positive, personal negative and personal positive consequences. This proved that the perception of appropriateness was dependent on the resulting consequences of the exchange of personal data.

5.2.4. PROVING INFORMATIONAL SUB-SPHERES

In Figures 5.10 and 5.11 you can find the relevant results printed from SPSS 22. In Figure 5.12 you can see the third conceptual model of this study, again with the results of the analysis included. And in Figure 5.13 these results are visually displayed. The absolute difference between the Appropriateness Of Intra-Sphere Personal Data Exchanges and the Appropriateness Of Inter-Sphere Personal Data Exchanges in this test was $\Delta\mu = 0.47531$. This was a perfect verification of the results of Figure 5.3 on page 78. Again, this result proved that the means are different with a significance of $p = 0.000$, although it did not give insight into the differences for different spheres.

The absolute difference between the Appropriateness Of Intra-Sub-Sphere Personal Data Exchanges and the Appropriateness Of Intra-Sphere Personal Data Exchanges was $\Delta\mu = 0.57099$. This was a large difference, considering that it was even larger than the difference between the perceived Appropriateness Of Intra-Sub-Sphere Personal Data Exchanges and Appropriateness Of Inter-Sub-Sphere Personal Data Exchanges. Again, this difference had a significance of $p = 0.000$. From this was concluded that hypothesis H15 was confirmed. Hypothesis H15 was defined as "The average appropriateness of personal data exchanges in sub-spheres is even higher (more appropriate) than the average appropriateness of personal data exchanges in normal spheres."

It should be noted that in the questions from the survey on which this analysis was based, the intra-sub-sphere questions were based on the sharing of personal data inside the respective organisation (e.g. inside the hospital or inside the university). In the theoretical chapter of this thesis other kinds of sub-spheres were discussed, for instance based on hierarchy or social relations. The existence of sub-spheres based on these terms were not proven by this research, solely the existence of sub-spheres created by organisations. It should also be noted that these results were only based on questions of the medical, educational and financial spheres.

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
SSAVGSUB – ISAVGSUB	,57099	1,16816	,11241	,34816	,79382	5,080	107	,000

Figure 5.10: Here you can find the results of the paired samples test between the intra-sub-sphere and intra-sphere mean appropriateness of data exchanges.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	ISAVGSUB – BSAVGSUB	,47531	,77366	,07445	,32773	,62289	6,385	107	,000

Figure 5.11: Here you can find the results of the paired samples test between the intra-sphere and inter-sphere mean appropriateness of data exchanges.

Figure 5.12: The third conceptual model developed for study 3 with the results included. This model overlapped with the model presented in Figure 5.2, but the main focus of this part of the research was in proving the existence of sub-spheres.

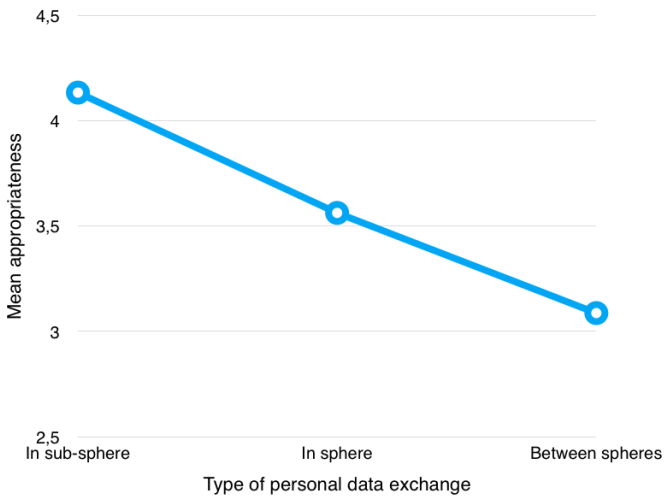


Figure 5.13: Here you can find the mean appropriateness of personal data exchanges in the same sub-sphere, in the same sphere and between different spheres. The survey questions on which these results were based, had an anecdotal formulation. This result proved the existence of sub-spheres for personal data exchanges and confirmed the result found in Figure 5.3.

5.3. EDX RELATED RESULTS

5.3.1. PERCEPTION OF PEOPLE ON THE PRIVACY CALCULUS

One of the topics integrated into the survey was the Perception Of People On The Privacy Calculus. The histogram of the survey item which best measured the degree of social acceptance of trading personal data for free services can be found in Figure 5.14. In the histogram can be seen that the opinions on the subject were very widespread, but the majority did accept that it is normal to trade personal data for free services in an educational MOOC setting. The average score of the question was $\Delta\mu = 4.46$. So apparently, in this sample the privacy calculus was largely accepted.

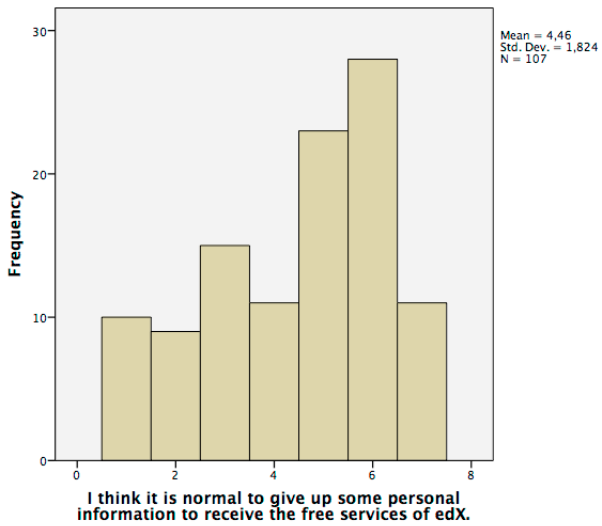


Figure 5.14: Here you can find the histogram for the statement "I think it is normal to give up some personal information to receive free services of edX." Respondents could indicate how much they agree or disagree with this statement by selecting an option on the seven point Likert scale anchored by "Strongly disagree" and "Strongly agree". This gave a representation of the Perception Of People On The Privacy Calculus.

5.3.2. APPROPRIATENESS OF DIFFERENT USES OF PERSONAL DATA OF EDX

In Figure 5.15 you can find the mean Appropriateness Of Different Uses Of Personal Data Of EdX. The uses are numbered in the graph and can be found in the list below. The different uses are sorted from most to least appropriate.

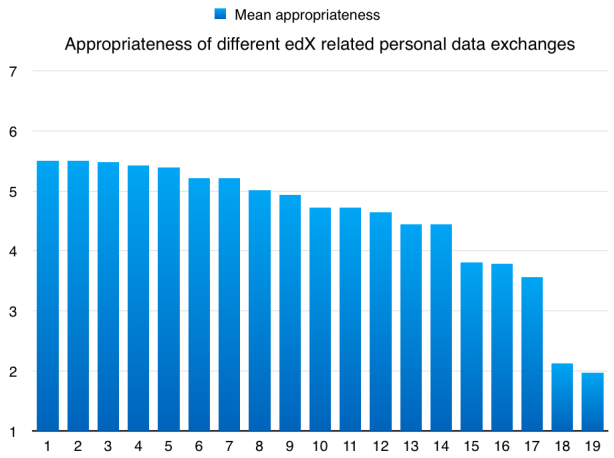


Figure 5.15: Here you can find the bar chart of the appropriateness of different edX related personal data uses. The uses of personal data are numbered and can be found back in the accompanying list.

To what extent do you find it appropriate for edX to use your edX data profile (i.e. survey response, demographics, performance, interaction, clicks) as follows? If edX uses your personal data to ...

1. To offer you new free courses
2. To advise you on better study methods
3. For anonymised scientific research
4. To create a comprehensive edX profile
5. To offer you job opportunities matching your interest and skills level
6. To connect you with mentors who may assist you
7. To select you for a masters study at a university in your vicinity
8. To offer you new paid courses
9. To suggest questions of other students to you that you may be able to answer
10. For anonymised market research
11. Connects your edX profile with educational data from other MOOC providers
12. To offer you educational offerings from third parties
13. To connect you with students that may be struggling
14. To connect you with paid mentors who may assist you
15. To offer you commercial offerings from third parties
16. Anonymise and openly share your data
17. Connects your edX profile with personal data from social network

18. Sells these profiles to companies

19. Sells your personal data to other organisations

A few things stood out in the bar chart presented in Figure 5.15. An expected drop in appropriateness was seen when the students' personal data is sold. Apart from this, the students also considered it slightly inappropriate when edX anonymises and openly shares their data, even though due of the anonymisation this would not cause them any discomfort.

5.3.3. INFLUENCE OF INTENTION

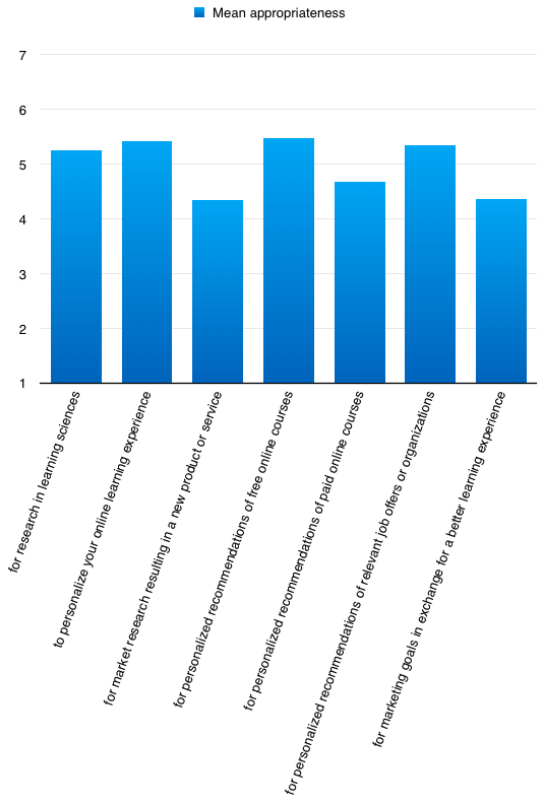


Figure 5.16: Here you can find the bar chart of the appropriateness of different edX related personal data uses with different intentions.

In Figure 5.16 you can see mean appropriateness for data uses of edX with different intentions. From this the Influence Of Intention could be deduced. It appeared that the intention of the data user did have some influence on the perceived appropriateness of the data use. What was remarkable was that the two intentions of data use that were related to market research and marketing goals, had the lowest mean appropriateness.

5.3.4. INFLUENCE OF DATA TYPE

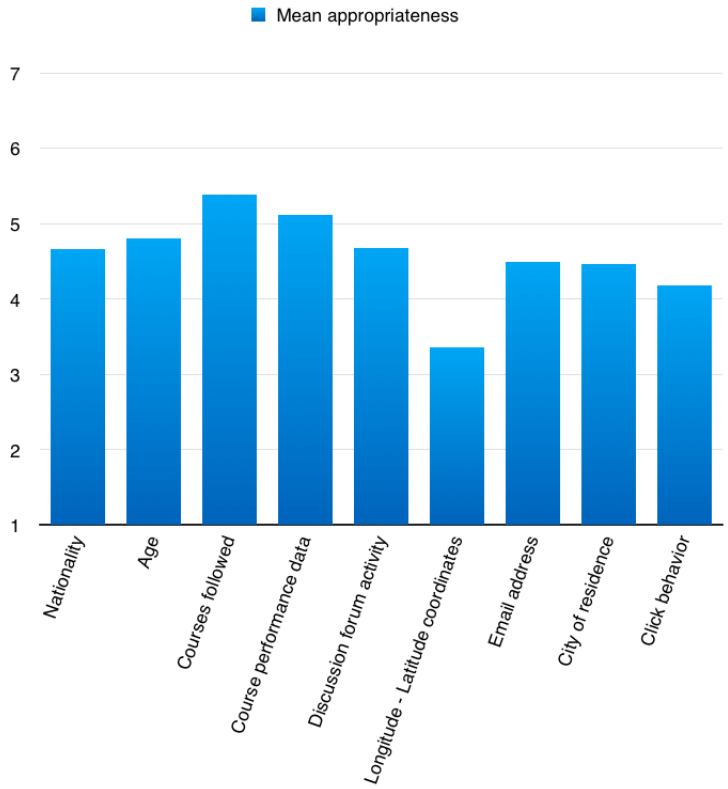


Figure 5.17: Here you can find the bar chart of the mean perceived appropriateness when edX uses different types of personal information.

In Figure 5.17 you can see the different mean perceived appropriatenesses when edX uses different types of personal information. From this the Influence Of Data Type could be deduced. What was interesting was that the students thought that using their exact clicking behaviour was not inappropriate. What also stood out here was the low perceived appropriateness of the exact longitude and latitude coordinates, it was to be expected that these data were perceived as sensitive.

This showed that edX users found it somewhat inappropriate for edX to collect and use their users' geographical location information. Ironically, edX does collect and use this information and shares this with researchers (for study 2 of this research for example). This raised the question whether it was morally just of me to use the latitude and longitude coordinates in this research. I think it was because of the following reason.

The question which was used to asses the perceived appropriateness of the respondents was "To what extent do you find it appropriate if edX collects and uses the following types of personal data about you?: Exact Longitude - Latitude coordinates". Because this concerned the exact location, people perceived this to be inappropriate. But, the latitude and longitude coordinates used in study 2 more or less assessed a respondent's approximate location. And as can be seen in Figure 5.17 the approximate location, for example an individual's city of residence, was perceived to be somewhere between "neutral" and "somewhat appropriate". So the use of the extracted latitude and longitude information in study 2 is considered morally just.

5.3.5. INFLUENCE OF DATA RECIPIENT

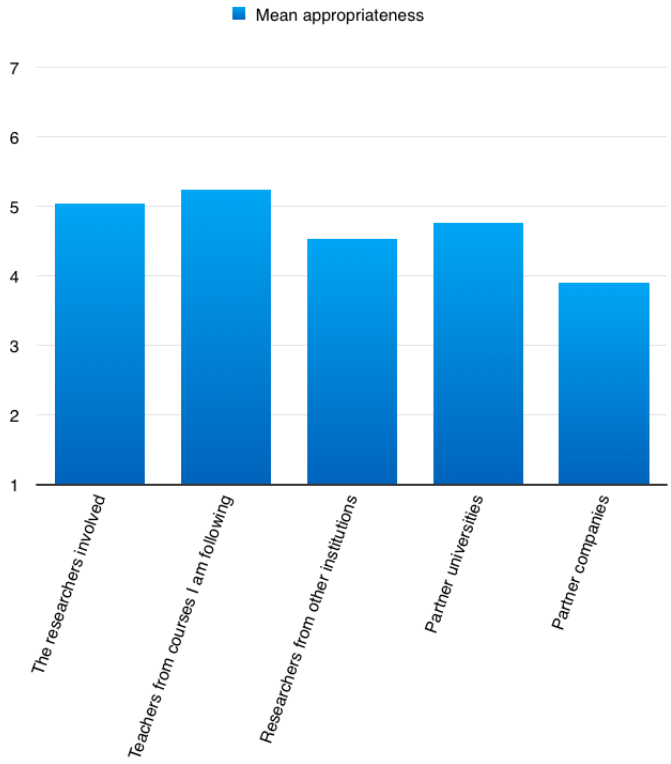


Figure 5.18: Here you can find the bar chart of the mean perceived appropriateness of edX related personal data exchanges with different data recipients.

In Figure 5.18 you can see mean appropriateness for exchanges of personal data by edX with different data recipients. From this the Influence Of Data Recipient could be deduced. Only data exchanges with partner companies were just below the neutral value of 4, so on average students found it slightly inappropriate when edX shares the personal data of its students with partner companies. The exchanges of personal data with other data recipients were not perceived as inappropriate.

5.4. SHORT SUMMARY OF RESULTS

The objectives of this study were to gain quantitative insight into the spheres of informational justice theory, quantitatively prove the influence of the consequence of a personal data exchange on its perceived appropriateness and prove the existence of sub-spheres. This was done by executing a quantitative survey based research. Evidence was found that the spheres of informational justice theory only applies to the governmental, medical and educational spheres and not to the commercial and financial spheres, indicating a clear discrepancy between the public and private domains. The salience of sub-spheres and the influence of the consequences of an exchange of personal data on its perceived appropriateness were also proven. In combination with the results of the previous studies this implied that especially private organisations should take steps to address the increasing Concerns For Information Privacy to mitigate any adverse effects.

OVERVIEW OF HYPOTHESES

No.	Hypothesis	Status
H1	Individual's overall concerns for information privacy have changed over the past 20 years.	Approved*
H2	Individuals with very little or very much knowledge of personal data collection and usage will have more concerns for information privacy than individuals with an average knowledge of personal data collection and usage.	Approved
H3	Individuals with higher levels of concern for information privacy will deem realistic alternatives to services which require the collection of personal information more important.	Approved
H4	The relation of hypothesis 3 will be affected by the frequency of use of services which require the collection of personal information.	Approved
H5A	Individuals with a higher level of education are more concerned for their information privacy.	Approved**
H5C	Individuals with a higher household income per household member are more concerned for their information privacy.	Approved
H5D	Individuals with a higher age are more concerned for their information privacy.	Approved
H5E	Individuals with a higher household income are more concerned for their information privacy.	Approved**/***
H9	Individuals with more concerns for information privacy will be less willing to disclose their information in a MOOC setting.	Rejected
H11A	Individuals who spend more time online will have less concerns for information privacy.	Rejected
H11B	Individuals who spend more time using social media will have less concerns for information privacy.	Rejected
H11C	Individuals who spend more time using online forums will have less concerns for information privacy.	Rejected
H11D	Individuals who use services which require the collection of personal data more often, will have more concerns for information privacy.	Approved
H10A	Financially independent individuals with more financial stress will have more concerns for information privacy.	Rejected
H10B	Financially independent individuals with more financial stability will have more concerns for information privacy.	Approved***
H10C	Individuals who are financially independent will have more concerns for information privacy than individuals who are financially dependent.	Approved
H5B	Individuals with a better position in society are more concerned for their information privacy.	Rejected
H12	Levels of concerns for information privacy will differ across countries.	Rejected
H13	The average appropriateness of personal data exchanges between spheres is lower (less appropriate) than the average appropriateness of personal data exchanges in the same sphere.	Approved****
H15	The average appropriateness of personal data exchanges in sub-spheres is even higher (more appropriate) than the average appropriateness of personal data exchanges in normal spheres.	Approved
H14	The average appropriateness of personal data exchanges is dependent on the consequence of the exchange, both in and between spheres. From least to most appropriate: personal negative consequence, societal positive consequence, personal positive consequence.	Approved

Table 5.2: Here you can see an overview of all hypotheses developed in all three studies and their results.

* This hypothesis is only approved for a field/student population in a organisational/online context in the US.

** These hypotheses were approved for a sub-sample of American and Canadian respondents.

*** These hypotheses were approved for a sub-sample of financially independent individuals.

**** This hypothesis was only proved for the medical, educational and governmental spheres.

6

DISCUSSION

The present research focuses on the antecedents of the CFIP construct. In the conducted studies a trend in the concerns for information privacy was found, as well as many, previously unexplored and significant relations between human values, demographics and the CFIP construct. The spheres of informational justice theory appears to only apply to the governmental, medical and educational spheres and the existence of sub-spheres and the influence of consequence have been proven.

6.1. SCIENTIFIC RELEVANCE

6.1.1. A TREND IN THE CONCERNS FOR INFORMATION PRIVACY

This research was the first attempt to date towards finding a trend of information privacy concerns. I have proven that concerns for information privacy were increasing, especially the concerns for unauthorized secondary use and improper access of personal information. This could give rise to many more research to elaborate on this trend. Of course, it is not the simple change in time that made people more concerned about their personal information, but other external factors.

Since this is the first time it has been proven quantitatively that the concerns for information privacy are increasing, this will have many implications. Previous literature focussed on how these concerns are established and created, but now research could focus on how these concerns can be reduced. Most people would agree that we do not want to live in a world where we are continuously concerned for our information privacy. The increase of these concerns was often speculated, but now that it is proven to exist, perhaps something can be done to safeguard the information privacy of society.

6.1.2. AN ADDITION TO MACKENZIE'S CERTAINTY TROUGH

In 1993, MacKenzie found a relation between the distance to the knowledge production and the degree of certainty in that knowledge. Over 2 decades later, I have proven that there exists a similar relation between the degree of knowledge of personal data collection and usage and the concern for information privacy. This newly found privacy concerns trough, similar to MacKenzie's theory, could be of much use in many future research in this field.

To further explain this relation, a complex conceptual model has been developed which also includes the constructs of "Importance of realistic alternatives", "Frequency of use of services which require the collection of personal data" and "Average time spend online". (Figure 4.38). This model is also considered to be a valuable scientific addition to the literature on the CFIP, because it gives a more thorough understanding of

the relation between the "knowledge of personal data collection and usage" and the concerns for information privacy.

6.1.3. REFINING THE INFORMATION PRIVACY CONCERNS RESEARCH FIELD

Study 2 confirmed many old and new relations between the CFIP construct and many other variables. For example, the relation between age and CFIP was confirmed. Some studies found a significant difference in the CFIP between men and women, but this discrepancy was, just as several other studies, not found in this study despite the large sample. This leads to the conclusion that the CFIP exists for men and women alike.

Apart from these confirmations, some new relations arose. A positive correlation was found between the highest level of education, the household income per household member and the financial stability of financially independent individuals on the one hand, and the concern for information privacy on the other hand. None of these correlations were found before, making these findings a valuable addition to the information privacy research field. The research of Milne *et al.* did find a significant relation between education and attitude towards privacy and an insignificant relation between income and attitude towards privacy in an Argentinian sample, but significant relations with the actual CFIP construct were not found in literature so far.

Also, by extending the concept of income to financial stress and financial stability, better insight was gained into the reason for the correlation between household income per household member and CFIP. It turns out that for financially independent individuals, financial stability causes more concerns for information privacy, since there is a significant positive correlation between financial stability and CFIP. And financial stress has no impact on ones concerns for information privacy, since there is no significant correlation between financial stress and CFIP. Not only are these newly found relations a valuable addition to the research field, but the newly developed scale could also be a helpful tool in different fields of research.

Another addition to this research field is an initial analysis of the influence of Schwartz' human values on CFIP. This found that people who cared more about universalism, self-direction, stimulation and security have more concerns for information privacy. This gives us more insight from a new perspective into the antecedents of the concern for information privacy. This new understanding could be used to find an improved explanation of the origin of individual's concerns for information privacy.

6.1.4. PROOF FOR THE SPHERES OF INFORMATIONAL JUSTICE

Study 3 is the first quantitative research into the spheres of informational justice theory. The results of this study indicate that intra-sphere exchanges of personal data are perceived as more appropriate than inter-sphere exchanges for the medical, educational and governmental spheres. The results also indicated that the consequence of an exchange of personal data influences its perceived appropriateness and sub-spheres have an important effect. Especially the clear distinction between the public and private domain is an interesting finding and is subject to further research. To be more precise: study 3 concluded that intra-sphere exchanges of personal data in the private domain (financial and commercial spheres) are considered even less appropriate than informational cross-contaminations between spheres in the private domain. This is considered a great scientific finding to which can be elaborated on in future research.

SUB-SPHERES

While some researchers refer to one informational sphere with several "sub-spheres" to indicate the different domains such as educational and medical domains, I see personal information as an additional good for the existing spheres and define sub-spheres as smaller bubbles inside these existing spheres such as educational and medical spheres. Having said that, the discovery of these sub-spheres could be a valuable addition to the literature in this new field of research. Similar to the discovery of the spheres of informational justice, this could also be the start of a new wave of research into this subject.

Also, the discovery of sub-spheres is a great way to combine the complementary theories of contextual integrity and spheres of informational justice. In this combination, the spheres of informational justice theory provides the explaining framework, in which different allocation criteria are defined and compared. While the contextual integrity theory focusses on the creation of contextual sub-spheres of different sizes inside this larger framework of spheres of informational justice. This line of reasoning should be investigated in future research.

6.2. PRACTICAL RELEVANCE

The retention information privacy has more and more become a topic of public debate, and developments of technology are increasingly perceived as a threat to people's information privacy. What was unclear until now was whether people's concerns for information privacy were actually increasing or not. The clarification of this simple fact has many practical implications which can be found below.

From the complex conceptual model developed in study 2 can be concluded that a need for realistic non-tracking alternatives has commerced. This is also backed-up by the finding from study 1 that concerns for information privacy are increasing, especially the concern for the unauthorized secondary use of personal information has increased. This has also many practical implications.

Same goes for the results of study 3. In particular the large discrepancy between the perceived appropriateness of exchanges of personal data of the private and the public spheres has large implications for private companies. Specific implications for management and policy can be found in the section below.

6.2.1. IMPLICATIONS FOR MANAGEMENT AND POLICY

As stated above, the knowledge that people's privacy concerns are increasing has many wide implications for the management of all companies handling personal information in some way. Nowadays, that includes almost every single company.

These increasing concerns for information privacy could create distrust and friction between a company and its customers or between companies and this could interfere with a company's objectives. To maintain good relations with customers as well as business partners, companies should take good care of the handling of personal information. To achieve this, management should adjust policy to ensure:

- the technical safety of the information storage system;
- a correct internal administration of access rights;
- the accuracy of the information;
- transparency and honesty towards the data owner about the use of the personal data.

When this is done and policy is adjusted correctly, these increasing concerns for information privacy could be limited and any adverse effects could be mitigated.

The empirical evidence of the increase of privacy concerns has even larger implications for corporations involved in social networks or social media. For these corporations, the issue of privacy is of even more significance. Therefore these companies should show even more respect for the privacy of its users and think of ways for its users to acquire increased control over their personal information. An example would be to add extra privacy respecting features to the social network.

From the complex conceptual model developed in study 2 can be concluded that a need for realistic non-tracking alternatives has commerced. This need could be addressed by creating new applications with more focus on the retention of information privacy or with additional privacy protecting features. Management

should change policy to reflect this shift in focus to privacy protecting applications and/or features, not only to prevent losing customers or partner companies, but to gain trust from customers and companies and perhaps even to gain revenue when pioneering in a specific market.

The significant positive correlations of education, financial stability and household income per household member with the CFIP construct found in study 2 have managerial implications for companies targeting wealthy, high educated people. Apparently this group has even higher privacy concerns and this should be taken into account when targeting this group. Similar to the previous policy recommendations, focussing on privacy protecting applications or extra privacy protecting features could be extra effective for this group, because of their elevated privacy concerns.

An important result of study 3 was the the large discrepancy between the perceived appropriateness of exchanges of personal data of organisations in the private and public sphere. As predicted, public companies should be extra careful when exchanging personal data with private companies. But unexpectedly, private companies should be even more careful when exchanging personal data with other private companies. This can be achieved by management by following the four point of advice which can be found in the above enumeration.

Another result of study 3 was the proof of the existence of organisational sub-spheres. In a way this contradicts the results of study 1, where it was stated that the concerns for improper access of personal information have been increasing. Apparently, exchanging people's personal data inside a company is only appropriate when the people inside the company receiving the information truly have previously specified access rights for that personal information. If this is the case, than this is another reason for management to focus extra on the correct management of the internal streams of personal data and the internal administration of access rights.

6.2.2. GOVERNMENTAL POLICY

Despite the increasing concerns for information privacy, many privacy infringing applications continue to enjoy large market shares.

This is because many applications have been locked into a dominant position by network effects (Katz and Shapiro, 1994). This mechanism allows dominant applications to do whatever they want without losing installed base, and this includes infringing people's privacy! Privacy protecting alternatives are almost always out-gunned by the already-settled applications and have no chance of competing.

The above stated problem could perhaps be solved by implementing national or even international governmental policy. Forcing large players into making their dominant applications compatible with privacy respecting alternatives would neutralize network effects and level the playground for privacy protecting alternatives.

6.2.3. IMPLICATIONS FOR EDX

The edX related research from study 3 encompassed many results which have implications for edX. The appropriateness of many specific data uses was measured as well as the influence of the intention of edX, data type and data recipient.

This analysis revealed that it is considered inappropriate by edX users for edX to sell student profiles to companies and to sell personal data to other organisations. Therefore, edX should try to refrain from doing this. None of the researched intentions were perceived as inappropriate by edX students. Same goes for different researched data recipients, the least appropriate data recipient was considered to be "partner companies", but this was only very slightly below neutral appropriateness. The analysis on the influence of the data type revealed that sharing someone's longitude latitude coordinates is considered somewhat inappropriate. Although the longitude latitude coordinates which edX extracts, (by using Qualtrics) are not very

accurate, edX should think about whether they should continue to share this information with researchers. Considering that this exchange of personal data is inside the educational sphere, people perceive this as neither appropriate nor inappropriate.

It is now known for a fact that privacy concerns have increased, especially concerns for the unauthorized secondary use and improper access of personal information. This has of course also implications for edX and for open and online education (O2E) in general. Organisations like edX do collect large amounts of data, so they should pay extra attention to preventing unauthorized secondary use and improper access of this data. Policy recommendations for achieving this are presented on page 93.

6.3. LIMITATIONS AND WEAKNESSES

One limitation of study 1 is that in some cases it was unclear what the actual date of the survey completion was, because this was sometimes not reported in the papers from which the data points were extracted. In the case of an unspecified survey completion date, the date had to be estimated. This was done by finding out when the paper was written (sometimes the paper was with the authors for revisions for years), and than subtracting another year to estimate the moment the surveys were completed. This approximation of one year is based on the other papers which did specify the date of survey completion. On average the first version of a paper is submitted to a journal one year after the survey completion.

A limitation of study 2 is that people's knowledge on personal data collection and usage activities is measured in a self-reporting manner. Although people have no reason to be dishonest, they are often mistaken. People can think they know everything even though they do not, and the other way around. Besides this, the question can be perceived differently. Knowing a few things about personal data collection and usage activities can be considered as much by some, but as little by others. It is therefore key to compare the results of this thesis with the results of previous research on this topic.

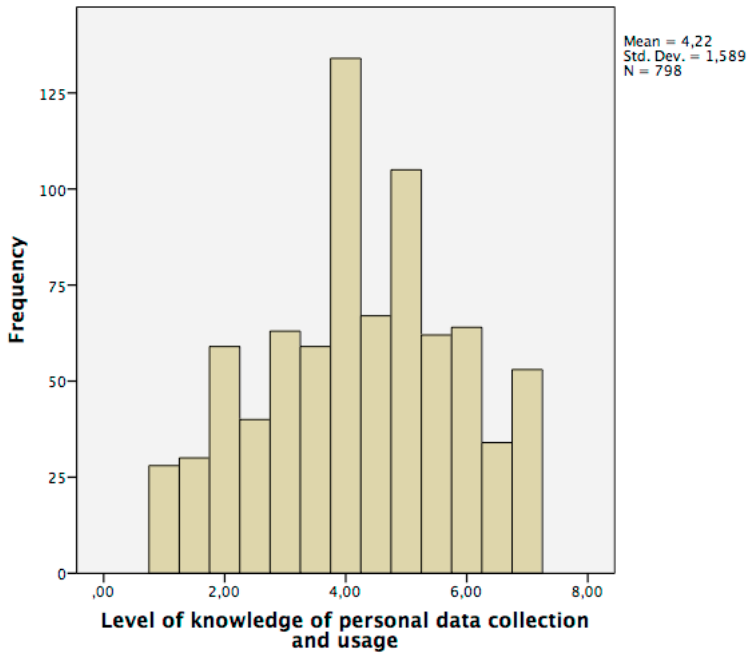


Figure 6.1: Here you can see the histogram of the degree of knowledge about personal data collection and usage from the data of study 2.

The most recent work on this topic is from 1992 (Nowak and Phelps, 1992). People's knowledge of information gathering and use was fairly low in 1992. More than twenty years ago, personal information was particularly used for marketing purposes and the concerns for information privacy were mostly related to

this. The research from Nowak and Phelps showed that the public had widely varying perceptions of the types of information accessible to marketers, as well as the accuracy of the collected personal information. When this is compared to the results from this research, which can be found in Figure 6.1, it can be seen that the perception of people's knowledge on personal data collection and usage activities is now also perceived to be widely varied and spread out around the neutral point. This is in accordance with the 22 years old research of Nowak and Phelps (1992).

This still does not prove that the self-reporting structure did not influence the results, but it is good to know that the spread in knowledge on personal data collection and usage activities matches previous results and can be intuitively understood. This also gives confidence in the precision of the results.

Another limitation of this research is that the results of study 2 on income could possibly be biased. These results are based on a sub-sample of Americans and Canadians, because the used question required a familiarisation with dollars. From the total of $N = 120$ only $N = 91$ respondents have indicated their income. Both hypotheses H5C and H5E are based on these responses, with H5E only using respondents who are financially independent, causing the results to be based on only $N = 64$ responses. This low sample size is not a problem, but the automated selection which cause the drop from $N = 120$ to $N = 91$ could have potentially caused a bias (e.g. if people with a higher income have less trouble sharing this information).

Another limitation of study 2 is the generalizability due to the high percentage of male respondents. This was caused by the MOOC in which the questionnaire was integrated, which was called AE 1110x "Introduction to aeronautical engineering". This technical field of interest is known to attract much more male than female participants, resulting in more male respondents for the integrated survey. For the main construct of CFIP there was only a very small difference between male and female respondents ($\Delta\mu = 0.0779$) which was insignificant ($p = 0.316$, equal variances not assumed). But the high percentage of male respondents could have had an indirect effect on the results. Also the sample has a higher than average percentage of well-educated people, due to the same reasoning. EdX in general attracts better educated people and on top of that the MOOC in which the questionnaire was integrated was technical, which is often perceived as difficult.

The same goes for study 3, this was a sub-sample of the sample used in study 2, so the sample of study 3 included the same high percentage of male respondents and high percentage of well-educated respondents.

6.4. FURTHER RESEARCH

The first study executed for this thesis had a retrospective character, in which existing data were used to find a trend in the concerns for information privacy. This was the only longitudinal study possible in the timespan of a master thesis, but setting up a true longitudinal study over one or two decades would allow for a much more specific research on the concerns for information privacy. For this research the CFIP construct was used, but this could be considered as an outdated construct. The recently developed internet privacy concerns construct (IPC) (Hong and Thong, 2013) should also be considered to use in this longitudinal study, although it does focus on internet concerns rather than organisational concerns. This construct is deemed fit, because it attempted to integrate all of literature's previously defined constructs of information privacy concerns and has proven to in general outperform them all. (Hong and Thong, 2013)

As previously discussed, another stream of research could find out why the concerns for information privacy have been changing over time. The influence of cultural changes, the rise of the internet and other technological developments on the concerns for information privacy could be investigated.

Unfortunately only five of the total of ten universal human values defined by Schwartz (1994) and their relation to the CFIP construct have been researched in this thesis due to a lack of space in the questionnaire. Further research could investigate the influence of the other five universal human values defined by Schwartz on CFIP: Power, achievement, benevolence, tradition and conformity.

As far as I am aware, the idea of informational sub-spheres inside the conventional information spheres

is new, so the discovery of this notion could unleash a new stream of research in this area. The research conducted in study 3 answered some questions, but raised even more. An example of an interesting addition to study 3 could be to try and find the existence of sub-spheres based on hierarchy or social relations. Study 3 only focused on the existence of sub-spheres based on a sub-sphere inside an organisation. Another idea for further research would be to link the differences in perceived appropriateness of different exchanges of personal data (in sub-sphere, in sphere or between spheres) to the willingness to disclose or even actual disclosure behaviour of these three different situations. Furthermore, new research could focus on using this new notion of sub-spheres as a method to link and integrate the theories of contextual integrity and spheres of informational justice with each other.

Another idea for further research lies in the clear distinction which was found between the public and private domain. Apparently people find it especially inappropriate if private companies exchange personal data with other private companies. This research has found this large gap in perceived appropriateness, but future research could dig into the reason behind this fact.

6.5. SUMMARY AND CONCLUSION

This study has concluded that on a global scale, people's concerns for unauthorized secondary use and improper access of their personal information has increased and in the US in a general/online context for a general/student population the overall concerns for information privacy have increased.

MacKenzie's certainty trough is found to also apply to the knowledge of personal data collection and usage activities and the concerns for information privacy, which proves the existence of a privacy concerns trough. A significant conceptual model has been developed to further explain this relation. Several relations between demographics and CFIP have been confirmed with respect to previous literature and for the first time, significant relations have been found between the highest level of education, financial stability and household income per household member and the CFIP construct. Four universal human values have also seemed to be of significance with respect to the CFIP.

It has been proven quantitatively that the spheres of informational justice theory only applies to the governmental, medical and educational spheres. The financial and commercial spheres very unexpectedly have an exact opposite result, in which intra-sphere exchanges of personal data are considered less appropriate than inter-sphere exchanges. Apart from this, the existence of sub-spheres and the influence of the consequence of an exchange of personal data have also been proved.

A

QUESTIONNAIRE STUDY 2

Below you can find the questionnaire which was used for the second study in this research. The complete questionnaire included several other questions about the course, such as motivations and expectations. These questions were not used in the research and are therefore left out of the survey below. The complete text below was presented to the respondents.

Welcome to edX and thanks for taking the survey! This survey will take about 5 to 10 minutes, and we will begin by asking you about your values, beliefs, and thoughts about the course. Remember there are no right or wrong answers. Just answer as accurately as possible. All of the information that you provide here is confidential, and will not be shared with the course instructors, or with other students. This information is used to help improve the course content in future years. Ultimately, your participation is voluntary, and you will not be penalized for not completing this section. The research is in accordance with the edX Privacy Policy, which can be read here. On behalf of the entire DelftX team, thank you very much.

Thieme, Pieter (Delft University of Technology),
Sasha (University of South Australia),
Omid and Phil (Stanford University)

Before you start filling out this survey, we would like to clarify two definitions which will be used throughout the survey . When we talk about personal information or personal data, we mean any information relating directly or indirectly to an identified or identifiable natural person. Common examples are your name, address, phone number, etc. When we talk about information privacy, we simply mean the privacy of personal information.

Here are some statements about personal information. From the standpoint of information privacy, please indicate the extent to which you agree or disagree with each statement by ticking the appropriate box. The seven boxes range from strongly disagree to strongly agree.

- It usually bothers me when companies ask me for personal information.
- All the personal information in computer databases should be double-checked for accuracy – no matter how much this costs.

- Companies should not use personal information for any purpose unless it has been authorized by the individuals who provided the information.
- Companies should devote more time and effort to preventing unauthorized access to personal information.
- When companies ask me for personal information, I sometimes think twice before providing it.
- Companies should take more steps to make sure that the personal information in their files is accurate.
- When people give personal information to a company for some reason, the company should never use the information for any other reason.
- Companies should have better procedures to correct errors in personal information.
- Computer databases that contain personal information should be protected from unauthorized access – no matter how much it costs.
- It bothers me to give personal information to so many companies.
- Companies should never sell the personal information in their computer databases to other companies.
- Companies should devote more time and effort to verifying the accuracy of the personal information in their databases.
- Companies should never share personal information with other companies unless it has been authorized by the individuals who provided the information.
- Companies should take more steps to make sure that unauthorized people cannot access personal information in their computers.
- I am concerned that companies are collecting too much personal information about me.
- I know which companies have personal data of me, what they are allowed to do with my personal data and also have a good understanding of the current (online) personal data collection activities of companies.
- I feel that I know enough about online personal data collection and usage to safely use services which require the collection of my personal data.
- I often use online services which require the collection of my personal data.
- I feel annoyed that I often have to use services which I rather would not, because of privacy reasons.
- Over the past ten years my attitude towards information privacy has become more and more tolerant.
- Compared to ten years ago, my attitude towards information privacy is now less tolerant.
- I would you be happy if edX researchers would use any additional available data of me to improve the quality of this research.

Here we briefly describe some people. Please read each description and think about how much each person is or is not like you. Tick the box to the right that shows how much the person in the description is like you. (Respondents were asked to indicate their answer by ticking one of the following boxes: Not like me at all, Not like me, A little like me, Somewhat like me, Like me, Very like me.)

- He strongly believes that people should care for nature. Looking after the environment is important to him.
- Thinking up new ideas and being creative is important to him. He likes to do things in his own original way.
- It is important to him to make his own decisions about what he does. He likes to be free to plan and to choose his activities for himself.

- He likes surprises and is always looking for adventures and new things to do. He thinks it is important to do lots of different things and to have an exciting life.
- Having fun and having a good time is important to him. He likes to "spoil" himself and do things that give him pleasure.
- It is important to him to live in secure surroundings and to be safe from threats from within and without. He avoids anything that might endanger his safety and is concerned that social order be protected.

Please respond to the following statements as accurately as you can. There are no right or wrong answers. Please be open and honest in your responding and indicate your answer by ticking one of the following boxes: Never, Rarely, Not very often, Moderately often, As often as not, Rather often, Extremely often.

- In an average day, how much do you feel you can reflect on the values that are important to you?
- On an average day, how much do you feel you can do the things that are most important to a sense of who you are?
- How often do you feel you get to step back and think about what is most important to you?
- How often are your relationships a source of frustration in your life?
- How often do you feel unable to complete everything you need to do?
- In an average day, how often do you feel you are falling behind?
- How often do things happen to you that make you feel you may be unable to keep up with life's hassles?
- How often do you receive positive feedback about your work?
- On an average day, to what extent do you feel competent in the things that matter to you?
- How often do you feel your skills and abilities are valued by others?
- In your life overall, how often do you find time to focus on what you care most about?
- How often do you feel a sense of accomplishment in your work?
- In your relationships, how often do you feel appreciated for who you are?

- How often do you use social media (i.e. Facebook, Twitter, Google+, Weibo, etc.)?
- How often do you contribute questions or answers to online forums?
- How much time do you averagely spend online in a day? (Please answer below in minutes)
- What is your nationality (country)?
- Where do you currently live?
- What is your ethnicity?
- How old are you?
- What is your marital status?
- Which of the following best describes your occupation
- What is the highest level of education that you have?
- In general, do you feel financially stable?
- In general, do you feel financial stress on a day-by-day basis?
- Are you financially independent? (By independent we mean you manage your own finances and do not depend on your parents)
- Think of this ladder as representing where people stand in your society. At the top of the ladder are people who are the best off – those who have the most money, the most education and the most respected jobs. At the bottom are the people who are the worst off – who have the least money, least education, and the least respected jobs or no job. The higher up you are on this ladder, the closer you are to the people at the very top; the lower you are, the closer you are to the people at the bottom. Where would you place yourself on this ladder? Please choose the number that indicates the rung where you think you stand at this time in your life, relative to other people in your society.
- Please indicate your total annual household income.

We are doing a lot of exciting research at Delft University of Technology. Do you want to be a pioneer and participate in one of our research projects? Let us know by clicking yes below, and we may send an email to you after the course. As a bonus, we will also include some fresh research results.

☐ Yes.
☐ No.

B

SURVEY INTEGRATION

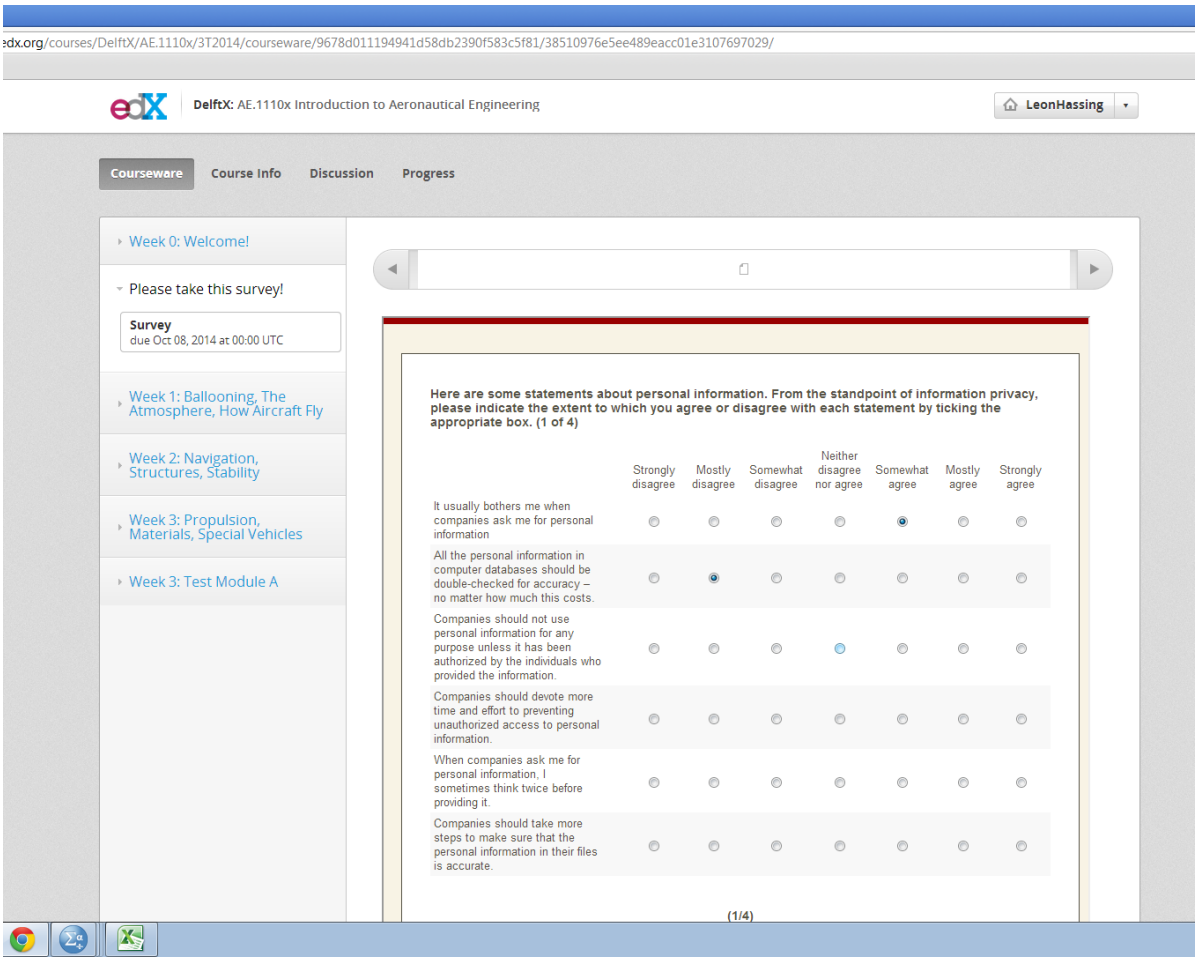


Figure B.1: Here you can see a screenshot of how a student, studying online for the course AE.1110x "Introduction to Aeronautical Engineering", could easily click on the survey button from the drop-down menu to the left, and start filling out the questionnaire.

C

LOCATION ACCURACY CHECK

Transform long-lat to countries

coords2country = function(

Help zoeken

```
1  library(sp)
2  library(rworldmap)
3  library(maptools)
4  library(maps)
5
6  points=read.csv2("Raw_lat-long_data_sorted_lat_ascending.csv")
7
8  coords2country = function(points)
9  {
10   countriesSP <- getMap(resolution='high')
11
12     pointsSP = SpatialPoints(points, proj4string=CRS(proj4string(countriesSP)))
13
14   indices = over(pointsSP, countriesSP)
15   indices$ADMIN
16 }
17
18 coords2country(points)
19
20 result<-coords2country(points)
21
22 write.csv(result, file = "location_accuracy_check.csv")
23
24
```

Figure C.1: Here you can see the code which was executed in R to transform the latitude and longitude information to their respective countries in a string format.

D

DATA TABLE STUDY 1

On the following pages you can find the data extracted from papers in literature which were used for the analyses in study 1.

Paper	Author	Journal	Publication year	Date of measurement	CFIP mean	Location	Population type	N	Context	Included?
Information Privacy in the Service sector: An Exploratory Study of Health Care and Banking Professionals	J.B. Earp and F.C. Payton	Journal of Organizational Computing and Electronic Commerce	2006	mei-01	5,02	US, southeast	Banking professionals	131	Banking	Yes
The Differing Privacy Concerns Regarding Exchanging Electronic Medical Records of Internet Users in Taiwan	H.G. Hwang, H.E. Han, K.M Kuo and C.F. Liu	Journal of Medical Systems Springer	2012	jan.-11	5,74	Taiwan	Members of an academic association related to healthcare information management	213	Electronic Medical Records	Yes
Situating the Concern for Information Privacy through an Empirical Study of Responses to Video Recording	D.H. Nguyen, A. Bedfor, A.G. Bretana, G.R. Hayes	Proceedings of the SIGCHI Conference on Human Factors in Computing Systems	2011	mei-10	5,81	US	Field	21	General	Yes
Opting-in or opting-out on the internet: does it really matter?	Yee-Lin Lai and Kai-Lung Hui	International Conference on Information Systems	2004	jan.-03	5,84	Singapore	Undergraduate and postgraduate students	32	General	Yes
An Empirical Investigation of Concerns of Everyday Tracking and Recording Technologies	D.H. Nguyen, A. Kobsa, G.R. Hayes	Proceedings of the 10th international conference on Ubiquitous computing	2008	sep.-07	6,00	US	Field	54	General	Yes
Relationship Marketing in Consumer Markets - A Comparison of Managerial and Consumer Attitudes about Information Privacy	A.J. Campbell	Journal of Interactive Marketing	1997	jan.-97	5,80	Canada, Toronto	Field	103	General	Yes
Users' perceptions on privacy and their intention to transact online: a study on Greek internet users	A. Zorotheos and E. Kafeza	Direct Marketing: An international journal	2009	sep.-08	5,76	Greece	Field	142	General	Yes
Measuring Individuals Concerns about Organizational Practices	Smith, Milberg & Burke	MISQ	1996	okt.-92	5,63	US, east	MBA's	146	General	Yes
Measuring Individuals Concerns about Organizational Practices	Smith, Milberg & Burke	MISQ	1996	apr.-93	5,56	US, east	Undergraduates	183	General	Yes
Effects of consumer characteristics on their acceptance of online shopping: Comparisons among different product types	J.W. Lian and T.M. Lin	Computers in human behavior	2008	jan.-07	5,89	Taiwan	Undergraduate students	216	General	Yes

The influence of personality traits and information privacy concerns on behavioral intentions	Korzaan, M. L., & Boswell, K. T.	Journal of computer information systems	2008	jan.-07	5,69	US	Undergraduate college students enrolled in an introductory level computer course	230	General	Yes
Measuring Individuals Concerns about Organizational Practices	Smith, Milberg & Burke	MISQ	1996	apr.-93	5,74	US	ISACA members	337	General	Yes
Predicting User Concerns About Online Privacy	M.Z. Yao, R.E. Rice and K. Wallis	JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY	2007	mrt.-05	5,63	US, south-west	Undergraduate students	413	General	Yes
An examination of the concern for information privacy in New Zealand regulatory context	E.A. Rose	Information & Management	2006	mrt.-04	5,80	New Zealand	Research staff, general staff and students	459	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,76	International	ISACA members	900	General	Yes
Study 2 of this thesis	L. Hassing	-	2014	sep.-14	5,80	International	Field	921	General	Yes
Strategies for reducing online privacy risks: why consumers read (or do not read) online privacy notices	G.R. Milne and M.J. Culnan	Journal of Interactive Marketing	2004	nov.-01	5,44	US	Faculty, staff and students at the authors' respective institutions	2468	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,83	Australia	ISACA members	100	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,95	Canada	ISACA members	100	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,65	Denmark	ISACA members	100	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,83	France	ISACA members	100	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,83	Japan	ISACA members	100	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,70	New Zealand	ISACA members	100	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,43	Thailand	ISACA members	100	General	Yes

Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,65	UK	ISACA members	100	General	Yes
Values, Personal Information Privacy, and Regulatory Approaches	S.J. Milberg, S.J. Burke, H.J. Smith and E.A. Kallman	Communications of the ACM	1995	jun.-95	5,73	US	ISACA members	100	General	Yes
Information Privacy in the Service sector: An Exploratory Study of Health Care and Banking Professionals	J.B. Earp and F.C. Payton	Journal of Organizational Computing and Electronic Commerce	2006	mei-01	5,26	US, southeast	Health care workers	163	Health care	Yes
Personality traits and concern for privacy: an empirical study in the context of location-based services	I.A. Junglas, N.A. Johnson and C. Spitzmuller	European journal of information systems	2008	aug.-07	5,63	US	Undergraduate and graduate business students	378	LBS	Yes
Understanding the Components of Information Privacy Threats for Location-Based Services	R.L. Raschke, A.S. Krishen and P. Kachroo	Journal of Information Systems	2014	jan.-13	5,57	US	Non-students	217	Nevada Department of Transportation (government)	Yes
Concern for Information Privacy and Online Consumer Purchasing	C. Van Slyke, J.T. Shim, R. Johnson and J. Jiang	Journal of the association for information systems	2006	aug.-04	5,95	US, southeast	Students	713	Online book purchase from Amazon.com	Yes
Concern for Information Privacy and Online Consumer Purchasing	C. Van Slyke, J.T. Shim, R. Johnson and J. Jiang	Journal of the association for information systems	2006	aug.-04	5,98	US, southeast	Students	287	Online book purchase from Half.com	Yes
International Differences in Information Privacy Concerns: A Global Survey of Consumers	S. Bellman, E. J. Johnson , S. J. Kobrin & G. L. Lohse	The Information Society: An International Journal	2004	jan.-02	5,80	International	Recruited online	140	Online environment	Yes
International Differences in Information Privacy Concerns: A Global Survey of Consumers	S. Bellman, E. J. Johnson , S. J. Kobrin & G. L. Lohse	The Information Society: An International Journal	2004	jan.-02	5,48	US	Recruited online	195	Online environment	Yes
Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model	N.K. Malhotra, S.S. Kim, J. Agarwal	Information Systems Research	2004	jan.-03	5,76	US	Field	449	Online environment	Yes
Adoption of Electronic Health Records in the Presence of Privacy Concerns	C.M Angst and R. Agarwal	MISQ	2009	okt.-08	5,68	US	Attendees of TEPR conference and online survey	366	Online health records	Yes

Table D.1: The listed papers in the above table have been used for the analyses in study 1.

E

PROCESS MACRO TEXT OUTPUT

The output text which resulted from computing the analysis for the standardized and centered model 17 of Hayes can be found below. The analysis below only includes age and gender as covariates.

```
Run MATRIX procedure:

************* PROCESS Procedure for SPSS Beta Release 120212 *************

       Written by Andrew F. Hayes, Ph.D. http://www.afhayes.com

**************************************************************************
Model = 17
    Y = ZCFIP
    X = ZIGNORME
    M = ZREALALT
    V = ZFREQ
    Q = ZONLINEC

Statistical Controls:
CONTROL= ZAGE   ZGENDER

Sample size
       798

**************************************************************************
Outcome: ZREALALT

Model Summary
         R       R-sq         F       df1       df2         p
     ,1744     ,0304    8,2999    3,0000  794,0000     ,0000

Model
             coeff        se         t         p
constant     ,0000     ,0349     ,0000    1,0000
ZIGNORME     ,1114     ,0352    3,1661     ,0016
ZAGE         ,1321     ,0352    3,7502     ,0002
ZGENDER     -,0573     ,0350   -1,6364     ,1021

**************************************************************************
Outcome: ZCFIP
```

```
Model Summary
         R       R-sq         F        df1        df2          p
     ,5539      ,3069    34,8412    10,0000   787,0000     ,0000

Model
             coeff         se          t          p
constant      ,0195      ,0304      ,6422      ,5209
ZREALALT      ,4272      ,0306    13,9525      ,0000
ZIGNORME      ,1047      ,0319     3,2771      ,0011
ZFREQ         ,1610      ,0313     5,1500      ,0000
ZONLINEC     -,0561      ,0307    -1,8256      ,0683
int\_1        -,0899      ,0256    -3,5059      ,0005
int\_2         ,0776      ,0453     1,7119      ,0873
int\_3        -,0258      ,0288     -,8971      ,3700
int\_4         ,0388      ,0213     1,8274      ,0680
ZAGE          ,1738      ,0303     5,7404      ,0000
ZGENDER       ,0693      ,0299     2,3188      ,0207

Interactions:

 int\_1    ZREALALT    X     ZFREQ
 int\_2    ZREALALT    X     ZONLINEC
 int\_3    ZIGNORME    X     ZFREQ
 int\_4    ZIGNORME    X     ZONLINEC

****************** DIRECT AND INDIRECT EFFECTS ************************

Conditional direct effect(s) of X on Y at values of the moderator(s)
      ZFREQ   ZONLINEC     Effect         SE          t          p
    -1,0000   -1,0000      ,0916      ,0513     1,7864      ,0744
    -1,0000     ,0000      ,1305      ,0482     2,7076      ,0069
    -1,0000    1,0000      ,1693      ,0540     3,1352      ,0018
      ,0000   -1,0000      ,0658      ,0386     1,7058      ,0884
      ,0000     ,0000      ,1047      ,0319     3,2771      ,0011
      ,0000    1,0000      ,1435      ,0381     3,7630      ,0002
     1,0000   -1,0000      ,0400      ,0448      ,8947      ,3712
     1,0000     ,0000      ,0789      ,0371     2,1287      ,0336
     1,0000    1,0000      ,1177      ,0406     2,9005      ,0038

Conditional indirect effect(s) of X on Y at values of the moderator(s)

Mediator
              ZFREQ   ZONLINEC     Effect   Boot SE   BootLLCI   BootULCI
ZREALALT    -1,0000   -1,0000      ,0490      ,0202      ,0148      ,0954
ZREALALT    -1,0000     ,0000      ,0576      ,0230      ,0161      ,1063
ZREALALT    -1,0000    1,0000      ,0663      ,0271      ,0183      ,1240
ZREALALT      ,0000   -1,0000      ,0390      ,0158      ,0121      ,0756
ZREALALT      ,0000     ,0000      ,0476      ,0182      ,0131      ,0848
ZREALALT      ,0000    1,0000      ,0562      ,0223      ,0155      ,1029
ZREALALT     1,0000   -1,0000      ,0289      ,0127      ,0093      ,0609
ZREALALT     1,0000     ,0000      ,0376      ,0143      ,0107      ,0669
ZREALALT     1,0000    1,0000      ,0462      ,0182      ,0125      ,0838

Values for quantitative moderators are the mean and plus/minus one SD from mean

****************** ANALYSIS NOTES AND WARNINGS ************************

Number of bootstrap samples for bias corrected bootstrap confidence intervals:
    10000
```

```
Level of confidence for all confidence intervals in output:
    95,00

------ END MATRIX -----
```

F

LISTS USED FOR STUDY 1

Three lists were used which assisted the search for usable data points in study 1. The references for these lists can be found below and on the following pages these lists are presented.

The list presented in Figure F.1 was extracted from: Bulgurcu, B. (2010). Antecedents and outcomes of information privacy concerns in online social networking: A theoretical perspective.

The list presented in Figures F.2 and F.3 was extracted from: Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: an interdisciplinary review. MIS quarterly, 35(4), 989-1016.

The list presented in Figures F.4 to F.10 was extracted from: Li, Y. (2011). Empirical studies on online information privacy concerns: literature review and an integrative framework. Communications of the Association for Information Systems, 28(1), 28.

Studies	Contexts	Outcome Variables
Angst and Agarwal 2009	E-Health	Opt-in Intention for e-health record use
Awad and Krishnan 2006	E-Com	Willingness to be profiled online for personalization
Chellappa and Sin 2005	E-Com	Likelihood for using personalized services
Culnan 1993	Direct marketing	Attitude toward secondary information use
Culnan and Armstrong 1999	Offline consumer transactions	The firm's attracting and retaining customers
Debatin et al. 2009	Online social networking	Change in privacy settings
Dinev and Hart 2006 Hann et al. 2007 Hui et al. 2007 Meinert et al. 2006 Malhotra et al. 2004	▪ E-Com ▪ Financial portals	Registering with a website Disclosure of personal information (willingness/intention)
Dinev and Hart 2005 Hine and Eve 1998 Pavlou et al. 2007 Phelps et al. 2001 Van Slyke et al. 2006	▪ E-Com ▪ Offline Commerce	Transaction (or purchase) intention
Miyazaki and Fernandez 2001	Online shopping	Willingness to pay for privacy Online purchasing rate
Egelman et al. 2009b	E-com	Willingness to examine multiple websites to find a better privacy protective option
Korzaan et al. 2009	Internet use	*Behavioural intentions* ▪ refuse to give information, ▪ take action to remove name, ▪ refuse to purchase
Lwin et al. 2007 Wirtz et al. 2007	Online Advertising	*Individual Responses* ▪ *Fabricate*: Misrepresentation of personal information ▪ *Protection*: Adoption of privacy protection technologies ▪ *Withhold*: Refusal to purchase from (or register to) a web site
Malhotra et al. 2004	▪ E-Com	▪ Trusting beliefs
Okazaki et al. 2009	▪ Mobile Advertising	▪ Risk beliefs
Pavlou et al. 2007	E-Com	▪ Perceived Uncertainty
Sheehan and Hoy 1999 Sheehan 2002	Online Advertising	▪ Notifying ISP about unsolicited e-mail ▪ Requesting removal from maligning list ▪ Flaming senders of unsolicited e-mail ▪ Registering for web sites ▪ Providing incomplete data during registration ▪ Providing inaccurate data during registration
Son and Kim 2008	Internet Use	▪ Refusal (information provision) ▪ Removal (private action) ▪ Negative word-of-mouth (private action) ▪ Complaining directly to online companies (public action) ▪ Complaining directly to 3rd party organizations (Public action)

Figure F.1: Here you can see the first list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table B3. Empirically Descriptive Statistics on General Privacy by Discipline

Discipline	Level	Methodology								
		Lab/Field Experiment	Survey	Case Study	Structured Interviews	Content Analysis	Mathematical Simulation	Law/Policy Analysis	Technology Description/ Analysis	Theory Building
MIS	Individual	Angst and Agarwal 2006; Berendt et al. 2005; Gideon et al. 2006; Hann et al. 2007. 2002; Hui et al. 2007; Liu et al. 2005; Lu et al. 2004; Xu 2007; Xu and Teo 2004; Xu et al. 2005	Ackerman et al. 1999; Awad and Krishnan 2006; Belanger et al. 2002; Berendt et al. 2005; Buchanan et al. 2007; Chellappa 2008; Chellappa and Sin 2005; Culnan 1993; Dinev et al. 2006, 2006b; Dinev and Hart 2004, 2006; Dinev et al. 2008; Earp and Baumer 2003; Earp and Payton 2006; Faja and Trimi 2006; Hahn 2002; Hoffman et al. 1999b; Hui et al. 2006; Jackson et al. 2003; Jensen et al. 2005; Lederer et al. 2003; Little et al. 2005; Malhotra et al. 2004; McKnight and Chervany 2001; Milberg et al. 1995; Moores 2005; Moores and Dhillon 2003; Pabl and Kobsa 2005; Ranganathan and Ganapathy 2002; Rensel et al. 2006; Sheehan 2002; Smith et al. 1996; Son and Kim 2008; Spiekermann et al. 2001; Stewart and Segars 2002; Suh and Han 2003; Tam et al. 2002; Van Slyke et al. 2006; Vijayasarathy 2004; Xu et al. 2003; Yao et al. 2007	Lankshear and Mason 2001	Hine and Eve 1998; Payton 2003	Earp et al. 2005; Ng-Kruelle et al. 2003	Chellappa and Shivendu 2008	Blanchette and Johnson 2002	Ackerman 2004; Bapna and Gangopadhyay 2006; Berkovsky et al. 2005, 2006; Canny 2002; Cingil et al. 2000; Cranor 1999; Cranor et al. 2006; Di Pietro and Mancini 2003; Enzmann and Schneider 2005; Gabber et al. 1997; Garfinkel et al. 2002; Hung et al. 2007; Ishitani et al. 2003; Jiang and Landay 2002; Juang et al. 2002; Karat et al. 2006; Lau et al. 1999; Lederer et al. 2004; Polat and Du 2005; Reagle and Cranor 1999; Shapira et al. 2005	
	Group									
	Organizational		Hoffman et al. 1999b; Milne and Culnan 2002; Peslak 2005	Bonner and Chiasson 2005; Culnan 2005; Karat et al. 2005; Paul et al. 1999; Siau and Kam 2006	Smith 1993	Anton et al. 2004; McRobb and Rogerson 2004; Schwaig et al. 2005; Schwaig et al. 2006	Thatcher and Clemons 2000		Adar et al. 2003; Hsu 2003; Li and Sarkar 2006	Adams and Blandford 2005; Li and Sarkar 2006
	Societal/Cultural/ International		Bellman et al. 2004; Dinev et al. 2006a, 2006b; Hinnant and O'Looney 2003; Milberg et al. 1995; Sagi et al. 2004; Shah et al. 2007; Zakaria et al. 2003							
Social Sciences	Individual	Fusilier and Hoyer 1980; Greenberg and Firestone 1977; Metzger 2007; Stone et al. 1983; Stone-Romero et al. 2003; Tolchinsky et al. 1981	Eoyang 1974; Marshall 1974; Metzger 2004; Pedersen 1997, 1999; Ribak and Turow 2003; Vinsel et al. 1980		Stone et al. 1983					
	Group									
	Organizational			Bonner 2007	Stanton and Stam 2003	Hong et al. 2005; LaRose and Rifon 2006				
	Societal/Cultural/ International		Davis and Silver 2004' Newell 1998		Altman 1977			Metzger and Docter 2003; Regal 2003; Taylor 2002		

Figure F.2: Here you can see the second list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table B3. Empirically Descriptive Statistics on General Privacy by Discipline (Continued)

Discipline	Level	Methodology								
		Lab/Field Experiment	Survey	Case Study	Structured Interviews	Content Analysis	Mathematical Simulation	Law/Policy Analysis	Technology Description/ Analysis	Theory Building
Law	Individual						Gould 1980	Beany 1966; Mizell 1998; Solove 2008; Westin 2001		
	Group									
	Organizational									
	Societal/Cultural/ International							Gavison 1980; Konvitz 1966; Nissenbaum 2004; Post 1989, 2000; Warren and Brandeis 1890		
Marketing	Individual	Milne 1977	Campbell 1997; Culnan 1995; Dommeyer and Gross 2003; LaRose and Rifon 2007; Lewin and Williams 2003; Milne and Boza 1999; Milne and Culnan 2004; Milne and Gordon 1993; Milne and Rohm 2000; Milne et al. 2004; Miyazaki and Krishnamurthy 2002; Miyazaki and Fernandez 2000, 2001; Norberg et al. 2007; Nowak and Phelps 1992; Pan and Zinkhan 2006; Phelps et al. 2001; Phelps et al. 2000; Rifon et al. 2005; Sheehan 1999; Sheehan and Hoy 1999, 2000; Xie et al. 2006				Rust et al. 2002			Norberg et al. 2007
	Group									
	Organizational		Culnan 2000; Dolnicar and Jordaan 2007			Miyazaki and Fernandez 2000; Phelps et al. 1994				
	Societal/Cultural/ International					Petrison and Wang 1995				
Management and Organization Science	Individual		Culnan and Armstrong 1999; Mael et al. 1996; Milberg et al. 2000; Sundstrom et al. 1980; Woodman et al. 1982							
	Group									
	Organizational	Eddy et al. 1999	Stone and Stone 1990						Kalvenes and Basu 2006	
	Societal/Cultural/ International							Smith 2001		
Economics and Finance	Individual	Huberman et al. 2005	Acquisti and Grossklags 2005a, 2005b				Acquisti2004; Calzolari and Pavan 2006; Huberman et al. 2005; Hui and Png 2005; Tang et al. 2005			
	Group									
	Organizational						Pendergast 2002	Shamis et al. 2005		
	Societal/Cultural/ International									

Figure F.3: Here you can see the second list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table A-1: Literature Reviewed in the Study					
Literature	**Research objective**	**Privacy construct and measurement**	**Research method**	**Subjects and sample size (N)**	**Data analysis method**
Andrade et al. [2002]	Examine approaches to encouraging self-disclosure of personal information on the Internet	Concerns for self-disclosure of information; New scale with three dimensions: concerns for identification information, sensitive information, and preferences and habits	Experiment	Undergraduate students in an U.S. university; N = 114	Analysis of variance (ANOVA)
Angst and Agarwal [2009]	Test the changes of individuals' attitudes and opt-in intentions in the adoption of electronic health records	Concerns for Information Privacy (CFIP); 2nd-order construct adapted from Smith et al. [1996]	Experiment	Individuals attending a health conference; N = 366	Structural equation modeling (SEM)
Ashley et al. [in press]	Study the factors that affect customer engagement in relationship marketing efforts	Privacy concerns; New scale	Survey	Households in the U.S.; N = 251	SEM
Awad and Krishnan [2006]	Examine the relationship between information transparency and willingness to be profiled online	Privacy concerns; 2 items adapted from past research	Survey	Internet users participating in product evaluations; N = 523	SEM
Bansal et al. [2008]	Examine how the quality of privacy policy statements and privacy assurance cues contribute to increased online trust	Privacy concerns; 2nd-order construct adapted from Smith et al. [1996]	Experiment	Students from an U.S. university; N = 674	SEM
Bansal et al. [2010]	Study the factors that affect an individual's intention to disclose health information online	Health information privacy concerns; 3 items adapted from past research	Experiment	Students from an U.S. university; N = 367	SEM
Bellman et al. [2004]	Examine international differences in information privacy concerns and the impact of three antecedents	CFIP; 2nd-order construct adapted from Smith et al. [1996]	Survey	Internet users from 38 countries; N = 534	Multivariate analysis of covariance (MANCOVA)
Buchanan et al. [2007]	Develop short Internet-administered scales measuring online privacy concern and behaviors (General Caution and Technical Protection)	Online privacy concerns; New scale with 16 unidimensional items	Survey	Students from an university in the U.K.; N1 = 515, N2 = 69, and N3 = 1,122	Factor analysis, correlations

Figure F.4: Here you can see the third list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table A-1: Literature Reviewed in the Study - Continued					
Casalo et al. [2007]	Analyze the influence of several factors on consumer trust in online banking	Perceived website privacy; 7 items adapted from past research	Survey	Spanish-speaking Internet users; N = 142	SEM
Cases et al. [2010]	Study the impact of several factors on email campaign effectiveness	Perceived privacy concerns; 3 items measuring perceived privacy	Survey	Shoppers of a Web company; N = 330	SEM
Cazier et al. [2008]	Study the factors that influence customers' intention to use radio frequency identification technologies (RFID)	Perceived privacy risk likelihood and perceived privacy risk harm; New scales	Survey	U.S. residents; N = 320	SEM
Chen et al. [2001b]	Investigate the relationship between consumer characteristics and online information privacy concerns	Information privacy concerns; 3 dimensions adapted from Smith et al. [1996]	Survey	Combination of students, faculty, and researchers in the U.S.; N = 340	Correlation
Chen et al. [2009]	Study the Privacy Concerns About Peer's Disclosure of one's information (PCAPD)	PCAPD; 7 items adapted from past research	Experiment	Students from a university; N = 209	Analysis of Covariance (ANCOVA) and regression
Chen and Rea [2004]	Study the factors that influence the use of privacy control techniques to protect personal information online	Privacy concerns; Two-dimensions adapted from Smith et al. [1996]	Survey	Undergraduate students; N = 102	Multiple regression
Cheung and Liao [2003]	Examine the supply-side hurdles in B2C e-commerce in Hong Kong	Privacy concerns; New scale	Survey	Hong Kong residents; N = 138	Multivariate regression
Chiu et al. [2009]	Understand e-shoppers' repurchase intentions	Privacy; 3 items adapted from past research	Survey	E-shoppers in Taiwan; N = 360	SEM
Cocosila et al. [2009]	Study the early investigation of new IT acceptance	Perceived privacy risks; 3 items adapted from past research	Experiment	Participants recruited from a university website; N = 303	SEM
Culnan and Armstrong [1999]	Study the impact of procedural fairness on the relationship between privacy concerns and customers' willingness to be profiled	Privacy concerns; New scale	Survey	U.S. households; N = 1,000	Discriminant analysis
Dai and Palvia [2009]	A comparative examination of factors affecting mobile commerce adoption	Privacy perception; New scale	Survey	A convenient sample of m-commerce users in China (N = 106) and students in the U.S. (N = 84)	SEM
Dinev et al. [2006]	Examines cross-cultural differences in beliefs related to e-commerce use for Italy and the United States	Privacy concerns; 4-item uni-dimensional construct adapted from Dinev and Hart, 2004, 2006	Survey	Individuals from Italy (N = 889) and the U.S. (N = 422)	SEM
Dinev and Hart [2004]	Develop an instrument to measure Internet privacy concerns and test the impact of two antecedents	Perceived privacy concerns; New scale with 2 dimensions: finding and abuse	Survey	Students and employees from universities and companies in the U.S.; N = 369	Regression
Dinev and Hart [2005]	Study the antecedents of privacy concerns and the intention to conduct online transactions	Internet privacy concerns; Information abuse dimension from Dinev and Hart [2004]	Survey	A combination of residents, teachers, students, and employees; N = 422	SEM

Figure F.5: Here you can see the third list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table A-1: Literature Reviewed in the Study - Continued					
Dinev and Hart [2006]	Study the impact of privacy risk beliefs on information privacy and the intention to provide personal information for online transactions	Internet privacy concerns; 4 items adapted from Smith et al. [1996] and Culnan and Armstrong [1999]	Survey	A combination of residents, teachers, students, and employees; N = 369	SEM
Dinev et al. [2008]	Test the relationship between Internet privacy concerns and the consequences under government surveillance	Internet privacy concerns; Two dimensions adapted from Dinev and Hart [2004, 2006]	Survey	A broad sample of individuals from various industries in the U.S.; N = 422	SEM
Eastlick et al. [2006]	Test the applicability of a traditional B2B relationship marketing framework to the B2C channel	Privacy concerns; 4 items adapted from focus group results and past research	Survey	U.S. households; N = 477	SEM
Faja and Trimi [2006]	Test the impact of a website's privacy interventions on users' perceptions and intentions during the initial interaction	General CFIP: adapted from Smith et al. [1996] and developed by authors; Perceived information privacy: adapted from past research and developed by authors	Experiment	Students from 2 U.S. universities; N = 210	ANCOVA and multiple regressions
Fogel and Nehmad [2009]	Study the associations between social networking user attributes and privacy concerns, risk taking and trust	Privacy concerns 3 items adapted from Dinev and Hart [2004]	Survey	College students in the U.S.; N = 205	ANOVA
Frye and Dornisch [2010]	Study the impact of topic intimacy and perceived privacy on the disclosure of information via instant messaging	Perceived privacy of a medium; Single item	Survey	Individuals from multiple nations; N = 214	Correlation and regression
Hoy and Milne [2010]	Examine gender differences in young adults' privacy beliefs, reactions to behavioral advertising and information sharing and privacy protection on social networks	Privacy concerns; Single item	Survey	Facebook.com users; N = 589	T-test
Hui et al. [2007]	Study the impact of privacy statements and privacy seals on information disclosure by individuals	Privacy concerns; Adapted from Smith et al. [1996]	Survey	Business students in Singapore; N = 109	Logistic regression
Janda [2008]	Study the impact of four consumer online concerns (privacy, security, etc.) on the likelihood of making online purchases, and the moderating role of gender	Privacy concerns; New scale	Survey	Nonstudent Internet users; N = 404	SEM
Janda and Fair [2004]	Identify eleven potential concerns people may have about the Internet, including privacy, fraud, etc.	Privacy concerns; New scale	Survey	Non-student Internet users; N = 440	T-test
Ji and Lieber [2010]	Study the link between personal identifiable information (PII) disclosure and privacy concerns	Worry about information disclosure online; Single item	Survey	Adult Internet users; N = 1,623	Logistic regression
Joinson et al. [2010]	Study the link between online privacy concerns and actual behavior	Privacy dispositions and perceived privacy; Adapted from past research	Survey and experiment	Students and Internet users from multiple nations; N1 = 759, N2 = 181	Correlation, ANOVA, and linear regression
Junglas et al. [2008]	Study the factors that influence CFIP	CFIP; 2nd-order construct adapted from Smith et al. [1996]	Survey	Undergraduate and graduate business students; N = 378	SEM

Figure F.6: Here you can see the third list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table A-1: Literature Reviewed in the Study - Continued					
Kim [2008]	Examine the impact of culture on trust determinants in computer-mediated transactions	Privacy concerns; New scale	Survey	Students from universities in the U.S. (N = 246) and South Korea (N = 199)	SEM
Kim et al. [2008a]	Test the impact of trust and risk in consumers' electronic commerce purchasing decisions	Perceived privacy protection; New scale	Quasi-experiment	Undergraduate students; N = 468	SEM
Kim et al. [2008b]	Examine the effects of an educational intervention on consumer's knowledge of security and privacy	Privacy concerns; 4 items adapted from past research	Quasi-experiment	Undergraduate students in an U.S. university; N = 125	t-tests, SEM
Korzaan and Boswell [2008]	Test the impact of personality traits on CFIP	CFIP; 2nd-order construct adapted from Smith et al. [1996]	Survey	Undergraduate students; N = 230	SEM
Krohn et al. [2002]	Study the potential influences of privacy concerns on consumers' attitudes toward websites and their satisfaction, etc.	Privacy concerns; Adapted from past research	Survey	College students from the U.S.; N = 219	Multiple regression
Kumar et al. [2008]	Investigate the factors that affect the use of security protection strategies by home computer users	CFIP; 2nd-order construct adapted from Stewart and Segars [2002]	Survey	Students from a public university in the U.S.; N = 120	SEM
Lai and Hui [2004]	Explain the differences in consumer participations in opt-in and opt-out configurations	Privacy concerns; 2nd-order construct adapted from Smith et al. [1996]	Experiment	Undergraduate and postgraduate students; N = 32	t-tests
Laric et al. [2009]	Study the impact of a number of factors on healthcare privacy concerns	Concerns for healthcare; information privacy New scale	Survey	MBA students from the U.S. and Canada; N = 225	ANOVA
Lee and Cranage [in press]	Study the effects of personalization and privacy assurance on customer responses to travel websites	Privacy concerns; Adapted from past research	Experiment	Undergraduate students in the U.S.; N = 120	ANOVA and regression
Li et al. [2009]	Examine how Web vendors may foster swift trust among customers	Perceived privacy; Adapted from past research	Experiment	College students; N = 224	SEM
Lian and Lin [2008]	Examine the effects of consumer characteristics (such as privacy concerns) on online shopping acceptance in the context of different products and services	Privacy concerns; Adapted from Smith et al. [1996]	Survey	Undergraduate students in Taiwan; N = 216	Regression
Liu et al. [2004]	Compare American and Taiwanese perceptions of online privacy and the impact on trust on websites	Perceived privacy; New scale	Experiment	Undergraduate and graduate students in the U.S. and Taiwan; N = 436	Correlation
Liu et al. [2005]	Study how perceived privacy relates to the behavioral intention to make an online transaction.	Perceived privacy; New scale	Experiment	Undergraduate and graduate students in the U.S.; N = 212	SEM
Luo and Seyedian [2003]	Test the moderating effects in contextual marketing and customer-oriented strategies	Privacy concerns; 5 items adapted from literature	Survey	Internet users in the U.S.; N = 180	Regression

Figure F.7: Here you can see the third list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table A-1: Literature Reviewed in the Study - Continued					
Lwin et al. [2007]	Test the mediating effect of privacy concern on the link between business policy and regulatory perceptions, and users' protective online responses	Online privacy concerns; Adapted from past research	Experiment	Adult Internet users from multiple nations; N1 = 180, N2 = 627	ANOVA
Malhotra et al. [2004]	Develop a new scale to measure Internet Users' Information Privacy Concerns (IUIPC)	IUIPC; New scale	Experiment	Household Internet users; N = 449	SEM
McCole et al. [2010]	Test the moderating effect of privacy and security concerns on the impact of trust on online purchasing attitudes	Privacy and security concerns; Adapted from past research	Survey	Employees in an New Zealand universities; N = 383	Hierarchical regression, ANOVA
Milberg et al. [2000]	Test the impact of regulatory approaches on information privacy, corporate management of personal data and consumer reactions	CFIP; 2nd-order construct adapted from Smith et al. [1996]	Survey	Members of a multi-national association; N = 595	SEM
Nam et al. [2006]	Study the factors that influence consumers' privacy concerns and their willingness to provide marketing-related personal information online	Privacy concerns; Adapted from past research	Survey	Internet users in Korea; N = 323	SEM
Okazaki et al. [2009]	Explores the consequences of consumers' privacy concerns in the mobile advertising context in Japan	Privacy concerns; Adapted from Malhotra et al. [2004]	Quasi-experiment	Japanese mobile users; N = 510	SEM
Pavlou et al. [2007]	Study the nature of online uncertainty and the mitigation approaches	Information privacy concerns; 6 items adapted from Smith et al. [1996] and other research	Survey	Visitors to an online bookstore (N1 = 268), and visitors to an prescription filling website (N2 = 253)	SEM
Phelps et al. [2001]	Examine the interrelationships among antecedents and consequences of privacy concerns	Privacy concerns; Single item	Survey	U.S. households; N = 556	Regression
Premazzi et al. [2010]	Study the roles of incentives and trust in customer information sharing with e-vendors	Privacy concerns; Adapted from Smith et al. [1996]	Experiment	Firm employees in Italy; N = 178	ANOVA, ANCOVA, and regression
Rensel et al. [2006]	Test people's willingness to use publicly-available computers for e-commerce transactions	Task privacy; Adapted from past research	Survey	Public library patrons in the U.S.; N = 137	SEM
Rifon et al. [2005]	Study the effects of Web privacy seals on trust and personal disclosures and the impact of several moderators such as privacy concerns	Privacy concerns; New scale	Experiment	Undergraduate students in the U.S.; N = 210	ANOVA
Roca et al. [2009]	Investigate how e-investors are influenced by perceived trust, security, privacy and other constructs	Perceived privacy; 4 items adapted from past research	Survey	Undergraduate students in Spain; N = 103	SEM
Rohm and Milne [2004]	Examine consumer concern regarding the collection and use of personal medical information	Privacy concerns regarding specific types of information; New scale	Survey	U.S. households; N = 1,508	z-test

Figure F.8: Here you can see the third list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table A-1: Literature Reviewed in the Study - Continued					
Sheehan [1999]	Investigate gender difference in online privacy concerns	Privacy concerns; New scale	Survey	U.S. households; N = 889	t-test
Sheehan and Hoy [1999]	Study online consumers' response to privacy concerns	Privacy concerns; New scale	Survey	Internet users in the U.S.; N = 889	Correlation
Sheng et al. [2008]	Examines how personalization and context can impact customers' privacy concerns and the intention to adopt ubiquitous commerce	Privacy concerns; 4 items adapted from Smith et al. [1996] and Dinev and Hart [2004]	Experiment	University students in the U.S.; N = 100	Regression
Shin [2010]	Test the effects of trust, security and privacy in social networking	Perceived privacy; Adapted from past research	Survey	College students in the U.S.; N = 323	SEM
Smith et al. [1996]	Develop an instrument to measure CFIP	CFIP; New, 4-dimensional scale	Survey	Multiple samples: business graduate students (N = 77), undergraduate students (N = 59; N = 87; and N = 83) in the U.S.	Regression and correlation
Son and Kim [2008]	Develop a taxonomy of information privacy-protective responses and to test the impact of some antecedents	Information privacy concerns; 4 items adapted from Dinev and Hart [2006]	Survey	Panel members of a market research firm; N = 523	SEM
Stewart and Segars [2002]	Examine the factor structure of the CFIP instrument by Smith et al. [1996]	CFIP; 2nd-order construct based on Smith et al. [1996]	Survey	U.S. consumers (mall-shoppers); N = 355	SEM
Stutzman et al. [2011]	Explore how privacy settings and privacy policy consumption affect the relationship between privacy attitudes and disclosure behaviors in Facebook.com	Privacy attitude; Adapted from past research	Survey	University students in the U.S.; N = 122	Logistic regression
Tsarenko and Tojib [2009]	Examine the driving factors of privacy concern	Privacy concerns; Adapted from Smith et al. [1996]	Survey	Australian consumers; N = 456	Hierarchical regression
Van Slyke et al. [2006]	Assess the impact of consumers' concerns for information privacy on their willingness to engage in online transactions	CFIP; 2nd-order formative construct adapted from Smith et al. [1996] and Stewart and Segars [2002]	Survey	Visitors to Amazon.com (N = 713) and to Half.com (N = 287) from the U.S.	SEM
Ward et al. [2005]	Examine online privacy concerns and willingness to provide financial and personal information	Privacy concerns; Single item	Experiment	University students in Australia; N = 315	ANCOVA
Wei et al. [2010]	Study the factors that influence users' behavioral responses to short message service (SMS) ads	Privacy concerns; New scale	Survey	College students in Singapore; N = 407	Hierarchical regression
Wirtz et al. [2007]	Study the causes and consequences of online privacy concerns	Privacy concerns; Adapted from past research	Survey	Adult Internet users; N = 182	SEM
Xu [2007]	Examine the factors that alleviate privacy concerns in mobile computing	Privacy concerns; 4 items adapted from Smith et al. [1996]	Experiment	Mobile phone users in Singapore; N = 179	SEM
Xu and Teo [2004]	Examine the factors that alleviate privacy concerns in mobile computing	Privacy concerns; 7 items adapted from Dinev and Hart [2004] and Smith et al. [1996]	Experiment	Undergraduate students in Singapore; N = 256	SEM

Figure F.9: Here you can see the third list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

Table A-1: Literature Reviewed in the Study - Continued					
Xu et al. [2008]	Examine the formation of individuals' privacy concerns	Privacy concerns; 5 items adapted from Smith et al. [1996]	Survey	Undergraduate and graduate students at three universities in the U.S.; N = 823	SEM
Yang and Wang [2009]	Test the impact of information sensitivity and compensation on privacy concern and behavioral intention	CFIP; 2nd-order construct adapted from Malhotra et al. [2004] and Smith et al. [1996]	Experiment	Students from 2 universities in China; N = 458	Multivariate regression and ANOVA
Yao et al. [2007]	Test the impact of a number of antecedents on information privacy concerns	Concerns about online privacy; 11 items adapted from Smith et al. [1996] were used to measure organizational privacy; these items, along with 9 additional items, were used to measure online privacy	Survey	Undergraduate students in an U.S. university; N = 413	SEM
Yao and Murphy [2007]	Explore voters' perceptions and intention to use remote electronic voting systems	Privacy; New scale	Survey	U.S. citizens; N = 453	SEM
Yao and Zhang [2008]	Study factors that predict users' online privacy concerns in Hong Kong	Privacy concerns; Adapted from Smith et al., 1996	Survey	Undergraduate students in Hong Kong; N = 332	SEM
Youn [2008]	Examine the impact of parental influence on teens' attitude toward privacy protection.	Teens' privacy concerns; Adapted from past research	Survey	Public high-school students in U.S.; N = 395	Regression
Youn [2009]	Study the determinants of online privacy concern and its influence on privacy protection behaviors among young adolescents	Privacy concerns; Single item	Survey	Middle school students in the U.S.; N = 144	Regression
Yousafzai et al. [2009]	Develop and validate a multi-dimensional model of trust for Internet banking	Perceived privacy; Adapted from past research	Survey	Internet banking users from the U.K.; N = 441	SEM
Zhang et al. [2002]	Compares the privacy concerns of online consumers in China and the U.S. and identify major factors related to these concerns	Privacy concerns; New scale	Survey	Students, faculty and managerial professionals from the U.S. (N = 340) and China (N = 106)	t-test
Zimmer et al. [2010]	Examine the link between intent to disclose information and the actual disclosure	Information privacy concerns; 6 items adapted from Malhotra et al. [2004]	Experiment	Business management students in the U.S.; N = 236	Regression analysis
Zviran [2008]	Study factors that affect online privacy concerns and how these concerns could affect the users' online behavior	Privacy concerns; 5 dimensions adapted from past research	Survey	Graduates from an Israeli university; N = 217	Pearson correlation and ANOVA

Figure F.10: Here you can see the third list of papers which was used to find as much papers with quantitative research to the CFIP construct as possible, to use in the analyses of study 1.

G

QUESTIONNAIRE STUDY 3

Below you can find the questionnaire which was used for the third study in this research. In all except one question the respondents were asked to give their answers on a seven point appropriateness Likert scale anchored by "Very inappropriate" and "Very appropriate". The complete text below was presented to the respondents.

Welcome to this survey about personal data exchanges. You have been contacted for this survey, because you indicated that you are interested in future TU Delft research in the pre-course survey of the edX MOOC "Introduction to Aeronautical Engineering". This survey contains 16 questions and will take you approximately 10 minutes to complete. Please take your time and answer each question accurately. The findings will be used for policy recommendations concerning personal data and will help us to get a better understanding of the global concerns about privacy. On behalf of the DelftX team and the TU Delft department of Ethics/Philosophy of Technology: Thank you very much!

Thieme
Leon

On the following pages there will be several descriptions of personal data transactions between different organisations. From an ethical point of view, please indicate the extent to which you find the exchange of personal data appropriate or inappropriate. You can choose from "Very inappropriate" to "Very appropriate" and everything in between.Please note that when we refer to "personal data" in the following questions we mean "information that can be used on its own or with other information to identify, contact, or locate a single person, including - but not restricted to - information such as your home address, your phone number, your email address or any other demographic data, user data or user generated-data"

☐ I have read and understand the above text.

To what extent do you find it appropriate when the following organisations, without your consent, would exchange your personal data, such as contact details and demographic data, with each other?

- If hospitals exchange your personal data with each other.
- If governmental bodies, such as tax offices or ministries, exchange your personal data with each other.

- If schools or universities exchange your personal data with each other.
- If commercial organisations, such as retail stores or large international companies, exchange your personal data with each other.
- If banks exchange your personal data with each other.

To what extent do you find it appropriate when a hospital, without your consent, would exchange your personal data, such as contact details and demographic data, with the following parties?

- With governmental bodies.
- With schools or universities.
- With retail stores or large international companies.
- With banks.

To what extent do you find it appropriate when a governmental body, without your consent, would exchange your personal data, such as contact details and demographic data, with the following parties?

- With hospitals.
- With universities or schools.
- With retail stores or large international companies.
- With banks.

To what extent do you find it appropriate when a university or school, without your consent, would exchange your personal data, such as contact details and demographic data, with the following parties?

- With hospitals.
- With government bodies.
- With retail stores or large international companies.
- With banks.

To what extent do you find it appropriate when a retail store or large international company, without your consent, would exchange your personal data, such as contact details and demographic data, with the following parties?

- With hospitals.
- With government bodies.
- With universities or schools.
- With banks.

To what extent do you find it appropriate when a bank, without your consent, would exchange your personal data, such as contact details and demographic data, with the following parties?

- With hospitals.

- With government bodies.
- With universities or schools.
- With retail stores or large international companies.

To what extent do you find it appropriate when a university, without your consent, would exchange your personal data, such as performance data or contact details, with the following parties in the following manner?

- With other universities for use in a research on learning practices.
- With commercial companies for use in a research on learning practices.
- With other universities to present you with new learning opportunities abroad.
- With relevant commercial companies to present you with job opportunities.
- With other universities and consequently you get overwhelmed with uninteresting folders.
- With commercial companies and consequently you get overwhelmed with uninteresting folders.

To what extent do you find it appropriate when a hospital, without your consent, would exchange your personal data, such as medical records or contact details, with the following parties in the following manner?

- With other hospitals for use in a research on new medicine.
- With commercial companies for use in a research on new medicine.
- With other hospitals and consequently they found a better treatment for your disease.
- With commercial companies and consequently they find a better treatment for your disease.
- With other hospitals and consequently these other hospitals publicly disclose your personal data.
- With insurance companies so they can change insurance policy fees according to individual risk factors.

To what extent do you find it appropriate when a university, without your consent, would exchange your personal data, such as grades, contact details, and demographic data, with the following parties?

- Within the university for use in a scientific research
- With another university for use in a scientific research
- With other private organisations for scientific research

To what extent do you find it appropriate when a hospital, without your consent, would exchange your personal data, such as medical records, contact details, and demographic data, with the following parties?

- Within the hospital for use in a scientific research
- With another hospital for use in a scientific research
- With other private organisations for scientific research

To what extent do you find it appropriate when a bank, without your consent, would exchange your personal data, such as financial records, contact details, and demographic data, with the following parties?

- Within the bank for use in a scientific research
- With another bank for use in a scientific research
- With other private organisations for scientific research

You are now almost to the end of the survey, only 5 questions remaining. The following questions will be about the personal data exchanges and uses related to edX. Please note that many questions are hypothetical and do not represent the actual activities of edX. The actual use (and restrictions to use) are described in the edX Privacy Policy, which can be found here.

To what extent do you find it appropriate for edX to use your edX data profile (i.e. survey response, demographics, performance, interaction, clicks) as follows?If edX...

- ..uses the anonymised data for market research
- ..uses the anonymised data for scientific research
- ...uses your personal data to offer you new free courses
- ...uses your personal data to offer you new paid courses
- ...uses your personal data to offer you educational offerings from third parties
- ...uses your personal data to offer you commercial offerings from third parties
- ..uses your personal data to offer you job opportunities matching your interest and skills level
- ...sells your personal data to other organisations
- ...uses your personal data to connect you with students that may be struggling
- ...uses your personal data to suggest questions of other students to you that you may be able to answer
- ...uses your personal data to connect you with mentors who may assist you
- ...uses your personal data to connect you with paid mentors who may assist you
- ...anonymises and openly shares your data
- ...uses your personal data to select you for a masters study at a university in your vicinity
- ...uses your performance data to advise you on better study methods
- ...uses your personal data from all courses you did on edX to create a comprehensive edX profile
- ...connects your edX profile with educational data from other MOOC providers (Coursera, FutureLearn, Udacity, etc.) to create detailed student profiles
- ...connects your edX profile with personal data from (social network) services (Facebook, Twitter, Google) to create detailed student profiles
- ...sells these profiles to companies

Personal data is more and more becoming a common currency. Although it has no precise value, it often occurs that people receive free services in exchange for their personal data (e.g. the Google search engine). Please indicate below how much you agree or disagree with the following statements. (The respondent was asked to answer these two questions according to a seven point Likert scale anchored by "Strongly disagree" and "Strongly agree".)

- I consider the personal data I have to share with edX as the price I pay for receiving free services.
- I think it is normal to give up some personal information to receive the free services of edX.

With the following statements, we would like to know about your concerns regarding the use or sharing of your personal data. To what extent do you find it appropriate if edX uses your personal data for the following?If edX uses your personal data...

- ... for research in learning sciences
- ... to personalize your online learning experienc
- ... for market research resulting in a new product or service
- ... for personalized recommendations of free online courses
- ... for personalized recommendations of paid online courses
- ... for personalized recommendations of relevant job offers or organisations
- ... for marketing goals in exchange for a better learning experienc

To what extent do you find it appropriate if edX collects and uses the following types of personal about you?

- Nationality
- City of residence
- Exact Longitude - Latitude coordinates
- Email address
- Age
- Courses followed
- Course performance data
- Discussion forum activity
- Click behaviour
- Course enrollment date and time

To what extent do you find it appropriate if the following parties have access to your edX profile and data?

- The researchers involved
- Teachers from courses I am following
- Researchers from other institutions
- Partner universities
- Partner companies

You are now at the end of this survey!

The data for this research will be explored by researchers at the TU Delft in the Netherlands and this research will likely encompass interesting results and possibly a publication. Would you like to receive the results and/or publication of this research?

☐ Yes, I would like to receive any results or publications.
☐ No, I would rather not receive any results or publications.

We are very grateful for your participation in this research. If, for some reason, you would rather not participate in any further research, please indicate this below.

☐ I would like to continue to help edX researchers and stay available for further research.
☐ I would not like to continue to help edX researchers and would rather not be contacted by edX researchers again.

H

REALISTIC ALTERNATIVES ANALYSIS

It was argued in the research proposal that one of the factors which encourages people to share their personal information, and therefore a factor which enlarges privacy infringement, is the lack of non-tracking alternatives for the most important web and telephone services. This section will give a clear overview of the current available non-tracking alternatives for search engines, internet browsers, online shopping sites, texting services, email services and telephone services.

H.1. SEARCH ENGINES

In Figure H.1 you can find the global market share distribution for search engines. As expected, the well-known players Google, Yahoo and Microsoft's search engine Bing all have large market shares, but the Chinese player Baidu claims a large piece of the pie. These four players together own 98.88 % of the entire global market share for search engines. Google, Yahoo and Bing all often track, store and sell your information to third parties, but whether the Chinese favourite search engine Baidu is also involved in these kind of activities is unclear. The company recently joined International Association of Privacy Professionals (Marketwired, March 2014), but then again, so did Google, Microsoft and Yahoo a long time ago.

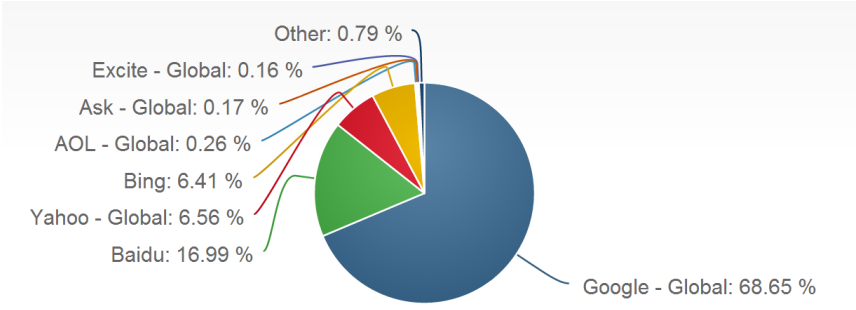


Figure H.1: Global market share statistics of search engine usage for all browsers, all operating systems for desktop type devices. (NETMARKETSHARE, April 2014)

Even though the large players dominate this market, small non-tracking alternatives are arising and growing. DuckDuckGo, Ixquick, Gibiru and Gigablast are all search engines which claim to be deliver the service of a search engine, but not violate privacy. There are also other alternatives: Blekko deletes its information after 48 hours, Ask.com has a function which will not store search history and hidemyass.com lets you use

our their proxy to surf anonymously online, hide your IP address, secure your internet connection, hide your internet history, and protect your online identity.

This spectrum of alternatives proves that people who truly have issues with information privacy have got the possibilities to receive the same services anonymously. So what makes people choose for the large internet search engines? This can be partly explained by the privacy paradox, which can be defined as the discrepancy between the statements of people's privacy concerns and their actual behaviour. Norberg, Horne and Home (2007) have found that for all information categories (personally identifying, financial, preferences, demographic, etc.), the level of actual disclosure significantly exceeded individual's intentions to disclose information. In other words, when push comes to shove, with the required service within reach, people do not set their privacy concerns at a high priority.

Other reasons for choosing the tracking search engines could be that most people do not know of these alternatives, or that people deem the search results of the popular search engines to be of higher quality and/or higher relevance.

H.2. INTERNET BROWSERS

In Figure H.2 you can find the global market share distribution for internet browsers.

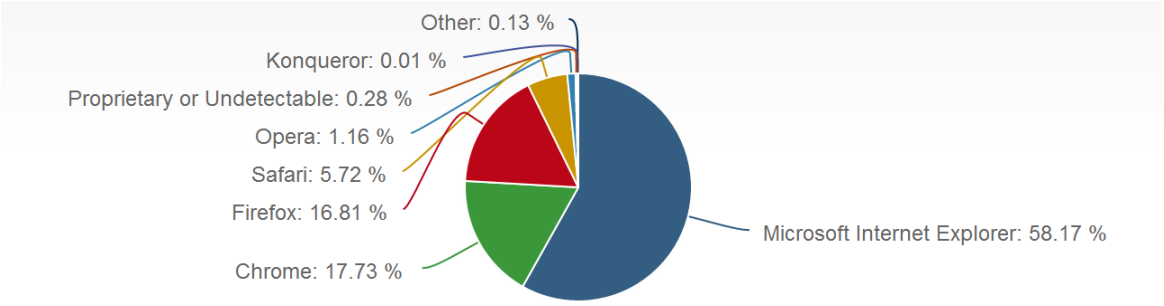


Figure H.2: Global market share statistics of internet browser usage for all operating systems for desktop type devices. (NETMARKETSHARE, May 2014)

When considering the level of privacy protection of internet browsers we have to distinguish two things. The first is the privacy intention of the browser, which is determined by the settings available and their default state. Second is the actual privacy, which can be compromised by technical flaws in the system which will allow hackers to collect your personal information.

In Figure H.3 you can find four settings related to privacy with the defaults of the four largest internet browsers from the browser security comparative analysis of NSS LABS (Abrams & Pathak, 2013). This will serve as an indicator for the privacy intention of the browsers. It should be noted that all browsers have incorporated the privacy protection mechanisms and we only consider the difference in default here. It can be seen that the largest browser, Microsoft internet explorer, has the highest privacy intention. The most remarkable differentiation of IE is the unique privacy feature called "Tracking Protection Lists", but the analysis of NSS LABS also states that "While the intent of the TPLs in IE is admirable, the current implementation makes certain add-ons, such as those provided by Abine and Disconnect, a superior choice for privacy.".

Product	Do Not Track	Third-Party Cookies	Geo Location	Tracking Protection List
Chrome	Not Set	Allow	Prompt	No
Firefox	Not Set	Allow	Prompt	No
Internet Explorer	On	Partial Block	Prompt	Built-In Option
Safari	Not Set	Block	Prompt	No

Figure H.3: Results of the browser security comparative analysis conducted by NSS LABS (Abrams & Pathak, 2013)

When we look at actual privacy protection, Microsoft Internet Explorer also seems to be one of the most secure browsers. In Figure H.4 you can see that vulnerability review conducted by Secunia reveals that Microsoft Internet Explorer has a relative small number of vulnerabilities and a low percentage of unpatched users.

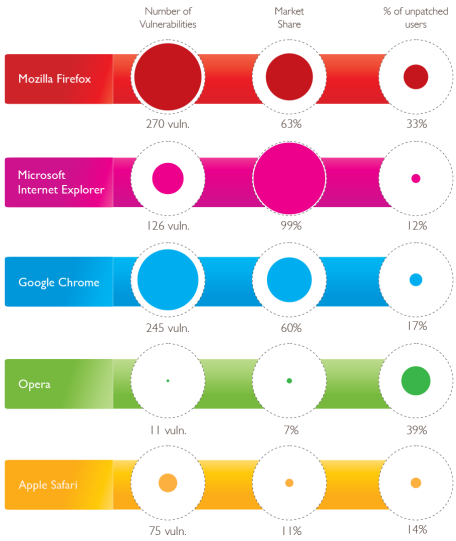


Figure H.4: The results of the Secunia Vulnerabiity Review 2014 (Secunia, 2014). Note that the market share figures differ from Figure H.2, since here the market share is based on instalment rather than use.

Although this sounds promising for Microsoft Internet Explorer, they have been experiencing some security flaws recently. On 27 April 2014 a bug was found in Microsoft Internet Explorer which could allow hackers as much access as a legitimate user (Campbell, 2014)

But all in all, the leading internet browsers all have taken steps to protect information privacy to some extent. Despite this, several internet browsers have appeared which focus even more on protecting privacy. Two examples are Comodo Dragon and WhiteHat Aviator. Comodo Dragon claims to have unsurpassed security and privacy features (Comodo, n.d.), and WhiteHat Aviator says they have the tightest security and privacy safeguards all built-in, all activated, all ready-to-go (WhiteHat Security, n.d.). So even if consumers feel their privacy is not protected well enough, there are alternatives which focus on privacy protection even more.

H.3. ONLINE SHOPPING SITES

When it comes to privacy protection of online web shops, it is always a clear choice for consumers. You either make use of the service and fill out your personal information (including address) or you do not fill out your information and you are unable to make use of the service. Another question that arises is what the company does with your information after you made use of the service.

The Dutch company Bol.com states in their privacy policy that they collect as much information about you as possible, including browsing behaviour and social medial, to optimize your shopping experience and personalise recommendations. What they do not do is personalise pricing or sell your personal data to third

parties. (Bol.com, n.d.)

The well-known amazon.com does all of the above too and goes even further regarding data collection and data sharing. When browsing Amazon via their app or the app of one of their subsidiaries, they collect information about your location, so they can offer location-based services. Also they share your personal information with several businesses, though they claim they are not in business of selling your personal information to others. (Amazon.com, n.d.)

Though some online shopping stores respect and protect privacy better than others. At this point, fully retaining ones privacy when buying products online seems to have no alternatives, with the exception of going to a physical store.

H.4. TEXTING SERVICES

The last few years, mobile messaging applications are overtaking the market of SMS services. Telecom companies usually retain records of all telephone conversations and SMS messages, in the US these telecom companies are obligated to do this by law. These records are called “Call detail records” and do not contain the contents of the phone call or SMS message, but other information like the phone numbers of the calling and called party and the call duration. However, the rise of the mobile messaging applications could introduce a step towards better privacy protection.

WhatsApp and Facebook Messenger are the largest global players, even though other players also take a piece of the pie. The privacy policy of WhatsApp states that they do in fact collect and use user provided information such as your mobile phone number, push notification name, billing information (if applicable) and mobile device information. They do not however retain the messages and pictures users send to each other. (WhatsApp Inc., 2012)

Facebook on the other hand saves all messages sent via either chat on a computer, or via Facebook Messenger. Despite this privacy breaching fact, Facebook Messenger is the second largest mobile messaging application of the entire globe.

So even though Facebook messenger is a large breach to people’s privacy, many people choose to stick with it, whilst there are countless alternatives. Of course, in this case the problem is very much related to network effects, a messaging service is only useful when your acquaintances also use that same service. That makes it even more astounding that people apparently use Facebook messenger instead of Whatsapp, since Whatsapp is the global leader by a long shot. In most countries in Europe more than half of the people use Whatsapp, but still Facebook Messenger has a significant installed base of 10 % - 30 %. But this will most likely change for the worse since Facebook has acquired Whatsapp in February 2014.

Even if one is not satisfied with the current privacy policy of Whatsapp, there are several privacy protecting alternatives. One example is Telegram, which uses an MTProto protocol to heavily encrypt data and even offers self-destructive messages. To prove their point, they held a contest and awarded 200.000 dollars to the first who could hack their application, but unfortunately nobody succeeded. (Telegram, 2014) Apparently the lack of realistic alternatives is not the reason which can explain the use of the privacy infringing messaging service of Facebook.

H.5. EMAIL SERVICES

In Figure H.5 you can find the market share distribution of email client usage in 2013.

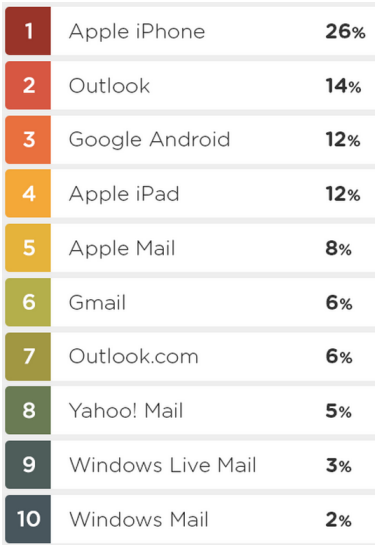


Figure H.5: Global market share statistics of E-mail client usage in 2013 based on the amount of emails opened in each client. (Smith, 2014)

The largest player in the top ten is Apple with 46 % of email opens. Fortunately Apple is very keen on privacy and they are known to respect their consumers' privacy. However, they do state in their terms of service that Apple is allowed to read "your account information and any content associated with that account". And also, people can only use this email service when they own an Apple product, which most people cannot afford.

Gmail does "open" your emails and uses automated processing of emails to provide personalised ads. Google even builds profiles of non-Gmail users who occasionally send emails to a Gmail account and use this information to target advertisements at you and your contacts. (Mick, 2013) Microsoft claims their Outlook.com offers more privacy than Gmail, but they use similar processing methods to find spam, grey mail, phishing scams, viruses, malware and other dangers and annoyances. (Zara, 2013)

So we can notice that all large email providing companies all use some sort of email processing, which can be seen as a privacy infringement. Fortunately realistic alternatives do exist. Hidemyass offers a receive-only email service which self-destructs in a pre-determined time which can be used to prevent revealing your real email address. Two other complete and free email service providers that focus on retaining your privacy are Vmail and Openmailbox. So there are definitely enough realistic alternatives if you are keen on retaining your privacy.

H.6. TELEPHONE SERVICES

As discussed before, most telecom companies are forced by governments to retain the call data records of all calls made. Also the EU member states are required to acquire and keep these call data records for six months to two years. (Leyden, 2005) In Figure H.6 you can find how long four American telecom providers retain their call data records as an example.

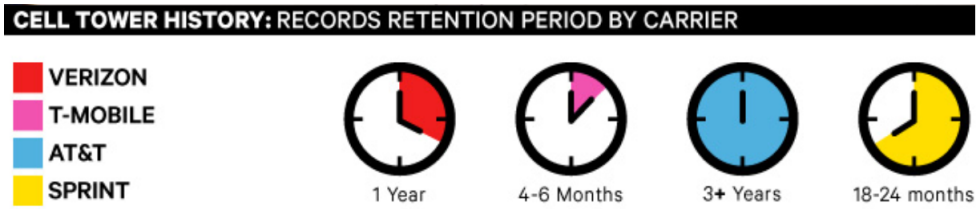


Figure H.6: Length of call data records retention of four different American telecom providers. (Kravets, 2011)

Since the retention of these files is obligated by governments in many countries, one would expect that there aren't any realistic privacy protection alternatives, but fortunately it seems so. February 2014 the Blackphone was announced, an Android phone which puts privacy first, and it is available since June 2014 (Souppouris, 2014). This phone offers peer-to-peer encrypted VoIP (Voice over Internet Protocol) calls and text messaging, a safeguarded contact list, 5 GB encrypted file storage, Wi-Fi manager to prevent Wi-Fi based tracking and a security centre which allows you to set app permissions system-wide or app-by-app. Needless to say this is a great privacy focussed solution, but the price tag of $ 629 could be a scare for some people.

So are any other (cheaper) alternatives for making private phone calls? There are many other VoIP apps like Skype, Viber and the iOS integrated Facetime, but are these VoIP services any safer than "normal" phone calls? Well, Skype does state in their privacy policy that they are allowed to store your instant messages, voicemails and video messages, but they need this permissions to provide you with the services they offer. But in July 2012 Skype has changed its peer-to-peer communications system to a server-based service, because their old "peer-to-peer encryption techniques would make it impossible for the company to comply with any government subpoenas for communications record" (Chaffin, 2013)

Most governments are catching up with technology and have laws and regulation which also apply to VoIP calls, making it impossible for VoIP services to be 100 % secure. Viber states that "if they receive a proper subpoena, they are obligated to provide records of who made and received calls, and when, but that no content from those conversations will be shared" (Garside, 2013).

So how can the Blackphone circumvent this obligation? Ben Dipietro (2014) states in his Wall Street Journal article that "The technology that allows Blackphone to operate without providing a backdoor for intelligence agencies is allowed because of an exemption in the Communications Assistance for Law Enforcement Act, Mr. Friedberg said.". I am not sure what to believe, but when it comes to making a 100 % secure phone call and circumventing organisations like the NSA, nothing seems to work. Especially since the CEO of Silent Circle himself states that "There is no such thing as an NSA-secure phone" (Souppouris, 2014).

So although retaining your privacy for 100% will be difficult when making a phone call, almost all telecom providers and other VoIP call providers do not store the contents of your phone calls or your messages and only disclose your call detail logs after an official court order.

H.7. SOCIAL NETWORK SITES

Even though social network sites are a platform in which people voluntarily share their data, I would still like to discuss a common privacy infringement accusation in this section.

One accusation to the world's most well-known social network site Facebook was that from 2006 to 2010 a slow shift took place in the default privacy settings, which would automatically share more and more personal information with the entire internet by default. (McKeon, n.d.) Personally I think this is a ridiculous accusation, since people who post their personal life on Facebook should be aware of which people they are sharing their information with. You cannot blame the corporation if you disagree with the defaults, if you do you should simply change them. But apparently Facebook does agree with the accusation (or just wants to work on their image) by resetting the defaults to "friends only". (Arce, 2014)

H.8. REFERENCES OF REALISTIC ALTERNATIVES ANALYSIS

Abrams, R., Pathak, J. (2013). BROWSER SECURITY COMPARATIVE ANALYSIS, Privacy Settings. Retrieved June 3, 2014 from https://www.nsslabs.com/reports/browser-security-comparative-analysis-privacy

Amazon.com (n.d.). Amazon.com Privacy Notice. Retrieved June 4, 2014 from http://www.amazon.com/gp/help/customer/display.html/ref=footer_privacy?ie=UTF8&nodeId=468496

Arce, N. (May 24, 2014). Finally, Facebook changes default privacy setting to “friends only” for new users. Retrieved June 12, 2014 from http://www.techtimes.com/articles/7493/20140524/finally-facebook-changes-default-privacy-setting-to-friends-only-for-new-users.htm

Bashir, A., (February 3, 2014). The Five Most Secure Web Browsers. Retrieved from http://www.therichest.com/business/technology/the-five-most-secure-web-browsers/4/

Bol.com (n.d.). Privacy Policy. Retrieved June 4, 2014 from http://www.bol.com/nl/m/voorwaarden/privacy-policy/index.html

Campbell, C. (April 28, 2014). Microsoft Admits to Huge Security Flaw in Internet Explorer. Retrieved June 3, 2014 from http://time.com/78828/internet-explorer-microsoft-security-flaw/

Chaffin, B. (July 11, 2013). The NSA Can Listen to Skype Calls (Thanks to Microsoft). Retrieved June 12, 2014 from http://www.macobserver.com/tmo/article/the-nsa-can-listen-to-skype-calls-thanks-to-microsoft

Comodo (n.d.). Comodo Dragon Internet Browser. Retrieved June 4, 2014 from http://www.comodo.com/home/browsers-toolbars/browser.php

Cutler, K.M. (December 4, 2012). The Reality Of The Global Messaging App Market: It’s Really Freaking Fragmented. Retrieved June 4, 2014 from http://techcrunch.com/2012/12/04/global-messaging-market/

Dipietro, B. (February 26, 2014). The Morning Risk Report: Blackphone Could Create a Black Hole for Compliance. Retrieved June 12, 2014 from http://blogs.wsj.com/riskandcompliance/2014/02/26/the-morning-risk-report-blackphone-could-create-a-black-hole-for-compliance/

Garside, J. (August 30, 2013). Viber founder: “People should be concerned about privacy”. Retrieved June 12, 2014 from http://www.theguardian.com/technology/2013/aug/30/viber-founder-talmon-marco-privacy

Kravets, D. (September 28, 2011). Which Telecoms Store Your Data the Longest? Secret Memo Tells All. Retrieved June 11, 2014 from http://www.wired.com/2011/09/cellular-customer-data/

Leyden, J. (December 14, 2005). MEPs vote for mandatory data retention. Retrieved June 11, 2014 from http://www.theregister.co.uk/2005/12/14/eu_data_retention_vote/

Marketwired (March 31, 2014). Baidu Joins International Association of Privacy Professionals, Reaffirms Global Commitment to User Privacy. Retrieved March 31, 2014 from http://www.marketwired.com/press-release/baidu-joins-international-association-privacy-professionals-reaffirms-global-commitment-nasdaq-bidu-1894129.htm

McKeon, M. (n.d.). The Evolution of Privacy on Facebook. Retrieved June 12, 2014 from http://mattmckeon.com/facebook-privacy/

Mick, J. (August 15, 2013). Google: Yes, we "Read" Your Gmail. Retrieved June 10, 2014 from http://www.dailytech.com/Google+Yes+we+Read+Your+Gmail/article33184.htm

NETMARKETSHARE, Market Share Statistics for Internet Technologies (April 2014). Desktop Search Engine Market Share. Retrieved May 28, 2014 from http://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustom=&qpcustomd=0

NETMARKETSHARE, Market Share Statistics for Internet Technologies (May 2014). Desktop Browser Market Share. Retrieved June 3, 2014 from http://www.netmarketshare.com/browser-market-share.aspx?qprid=0&qpcustomd=0

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. Retrieved June 2, 2014 from

Secunia (February 26, 2014). Secunia Vulnerability Review, Key figures and facts from a global IT-Security perspective. Retrieved June 4, 2014 from http://secunia.com/vulnerability-review/browser_security.html

Smith, L (January 16, 2014). Email Client Market Share: Where People Opened in 2013. Retrieved June 9, 2014 from https://litmus.com/blog/email-client-market-share-where-people-opened-in-2013/litmus-email-client-market-share-2013-infographic

Souppouris, A. (February 24, 2014). Blackphone: an Android phone that puts privacy first. Retrieved June 11, 2014 from http://www.theverge.com/2014/2/24/5441642/blackphone-silent-circle-geeksphone-pre-order-launch

Telegram (2014). $ 200.000,- to the hacker who can break Telegram. Retrieved June 9, 2014 from https://telegram.org/crypto_contest

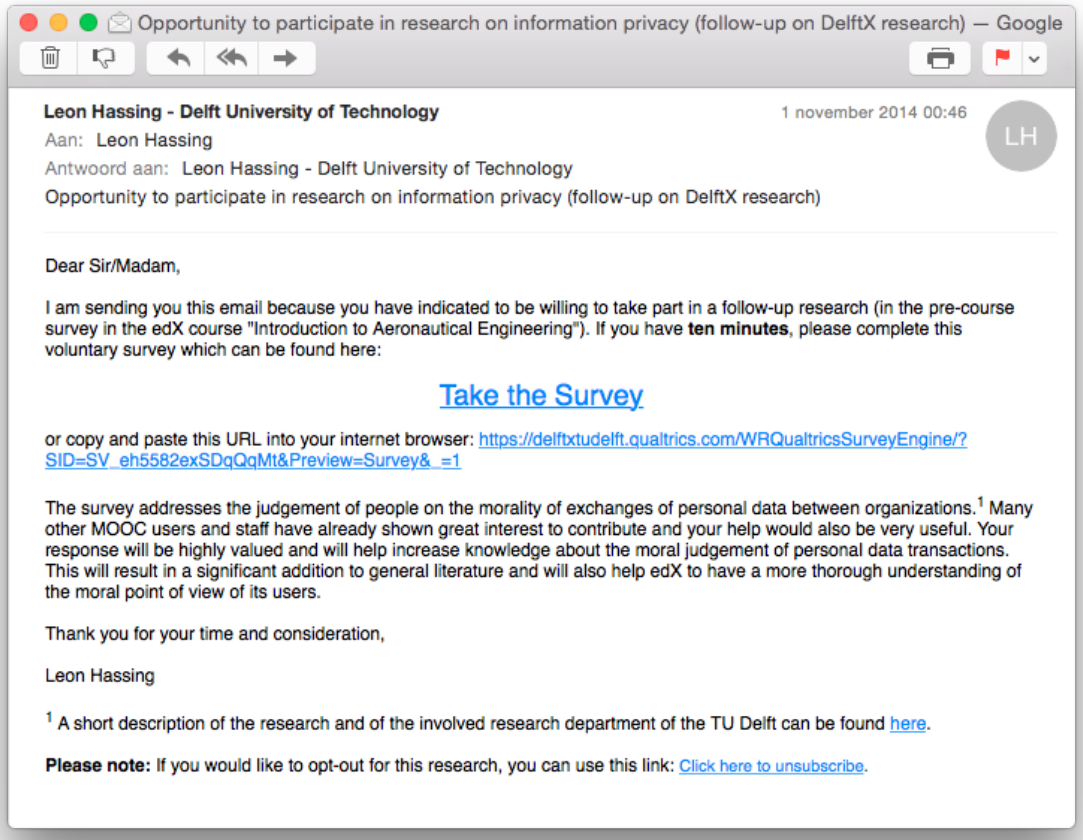
WhatsApp Inc. (July 7, 2012). WhatsApp Legal Info. Retrieved June 4, 2014 from http://www.whatsapp.com/legal/?l=en

WhiteHat Security (n.d.). The WhiteHat Aviator Web Browser. Retrieved June 4, 2014 from https://www.whitehatsec.com/aviator/

Zara, C. (February 16, 2013). Microsoft Rips Email-Snooping Google, But Is Outlook Any More Private Than Gmail? Retrieved June 10, 2014 from http://www.ibtimes.com/microsoft-rips-email-snooping-google-outlook-any-more-private-gmail-1094118

I

INVITATION EMAIL STUDY 3



Opportunity to participate in research on information privacy (follow-up on DelftX research) — Google

Leon Hassing - Delft University of Technology 1 november 2014 00:46

Aan: Leon Hassing

Antwoord aan: Leon Hassing - Delft University of Technology

Opportunity to participate in research on information privacy (follow-up on DelftX research)

LH

Dear Sir/Madam,

I am sending you this email because you have indicated to be willing to take part in a follow-up research (in the pre-course survey in the edX course "Introduction to Aeronautical Engineering"). If you have **ten minutes**, please complete this voluntary survey which can be found here:

Take the Survey

or copy and paste this URL into your internet browser: https://delftxtudelft.qualtrics.com/WRQualtricsSurveyEngine/?SID=SV_eh5582exSDqQqMt&Preview=Survey&_=1

The survey addresses the judgement of people on the morality of exchanges of personal data between organizations.[1] Many other MOOC users and staff have already shown great interest to contribute and your help would also be very useful. Your response will be highly valued and will help increase knowledge about the moral judgement of personal data transactions. This will result in a significant addition to general literature and will also help edX to have a more thorough understanding of the moral point of view of its users.

Thank you for your time and consideration,

Leon Hassing

[1] A short description of the research and of the involved research department of the TU Delft can be found here.

Please note: If you would like to opt-out for this research, you can use this link: Click here to unsubscribe.

Figure I.1: Here you can see the invitation email which was send to 532 potential participants of study 3.

BIBLIOGRAPHY

[1] J. Rauhofer, *Privacy is dead, get over it! 1 information privacy and the dream of a risk-free society,* Information & Communications Technology Law **17**, 185 (2008).

[2] N. M. Richards and J. H. King, *Big data and the future for privacy,* Available at SSRN (2014).

[3] H. J. Smith, S. J. Milberg, and S. J. Burke, *Information privacy: measuring individuals' concerns about organizational practices,* MIS quarterly , 167 (1996).

[4] K. A. Stewart and A. H. Segars, *An empirical examination of the concern for information privacy instrument,* Information Systems Research **13**, 36 (2002).

[5] N. K. Malhotra, S. S. Kim, and J. Agarwal, *Internet users' information privacy concerns (iuipc): the construct, the scale, and a causal model,* Information Systems Research **15**, 336 (2004).

[6] T. Dinev and P. Hart, *Internet privacy concerns and their antecedents-measurement validity and a regression model,* Behaviour & Information Technology **23**, 413 (2004).

[7] T. Buchanan, C. Paine, A. N. Joinson, and U.-D. Reips, *Development of measures of online privacy concern and protection for use on the internet,* Journal of the American Society for Information Science and Technology **58**, 157 (2007).

[8] S. Janda and L. L. Fair, *Exploring consumer concerns related to the internet,* Journal of Internet commerce **3**, 1 (2004).

[9] A. N. Joinson, U.-D. Reips, T. Buchanan, and C. B. P. Schofield, *Privacy, trust, and self-disclosure online,* Human–Computer Interaction **25**, 1 (2010).

[10] M. Laric, D. Pitta, and L. Katsanis, *Consumer concerns for healthcare information privacy: a comparison of us and canadian perspectives,* Research in Healthcare Financial Management **12**, 93 (2009).

[11] M. L. Korzaan and K. T. Boswell, *The influence of personality traits and information privacy concerns on behavioral intentions,* Journal of Computer Information Systems **48**, 15 (2008).

[12] I. A. Junglas, N. A. Johnson, and C. Spitzmüller, *Personality traits and concern for privacy: an empirical study in the context of location-based services,* European Journal of Information Systems **17**, 387 (2008).

[13] Y. Li, *Empirical studies on online information privacy concerns: literature review and an integrative framework,* Communications of the Association for Information Systems **28**, 28 (2011).

[14] M. Van den Hoven, *Privacy or informational injustice,* Ethics and electronic information in the 21st century , 139 (1999).

[15] H. Nissenbaum, *Privacy as contextual integrity,* Wash. L. Rev. **79**, 119 (2004).

[16] S. J. Milberg, S. J. Burke, H. J. Smith, and E. A. Kallman, *Values, personal information privacy, and regulatory approaches,* Communications of the ACM **38**, 65 (1995).

[17] M. Korzaan, N. Brooks, and T. Greer, *Demystifying personality and privacy: An empirical investigation into antecedents of concerns for information privacy,* Journal of Behavioral Studies in Business **1**, 1 (2009).

[18] N. Mohamed and I. H. Ahmad, *Information privacy concerns, antecedents and privacy measure use in social networking sites: Evidence from malaysia,* Computers in Human Behavior **28**, 2366 (2012).

[19] S. H. Schwartz, *Are there universal aspects in the structure and contents of human values?* Journal of social issues **50**, 19 (1994).

[20] J. Dawes, *Do data characteristics change according to the number of scale points used,* International Journal of Market Research **50**, 61 (2008).

[21] D. A. MacKenzie, *Inventing accuracy: A historical sociology of nuclear missile guidance* (MIT press, 1993).

[22] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2008), ISBN 3-900051-07-0.

[23] A. F. Hayes, *Process: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling,* Manuscript submitted for publication (2012).

[24] J. B. Earp and F. C. Payton, *Information privacy in the service sector: An exploratory study of health care and banking professionals,* Journal of Organizational Computing and Electronic Commerce **16**, 105 (2006).

[25] H.-G. Hwang, H.-E. Han, K.-M. Kuo, and C.-F. Liu, *The differing privacy concerns regarding exchanging electronic medical records of internet users in taiwan,* Journal of medical systems **36**, 3783 (2012).

[26] D. H. Nguyen, A. Bedford, A. G. Bretana, and G. R. Hayes, *Situating the concern for information privacy through an empirical study of responses to video recording,* in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, 2011) pp. 3207–3216.

[27] Y.-L. Lai and K.-L. Hui, *Opting-in or opting-out on the internet: Does it really matter?* ICIS 2004 Proceedings (2004).

[28] D. H. Nguyen, A. Kobsa, and G. R. Hayes, *An empirical investigation of concerns of everyday tracking and recording technologies,* in *Proceedings of the 10th international conference on Ubiquitous computing* (ACM, 2008) pp. 182–191.

[29] A. J. Campbell, *Relationship marketing in consumer markets: A comparison of managerial and consumer attitudes about information privacy,* Journal of Interactive Marketing **11**, 44 (1997).

[30] A. Zorotheos and E. Kafeza, *Users' perceptions on privacy and their intention to transact online: a study on greek internet users,* Direct Marketing: An International Journal **3**, 139 (2009).

[31] J.-W. Lian and T.-M. Lin, *Effects of consumer characteristics on their acceptance of online shopping: Comparisons among different product types,* Computers in Human Behavior **24**, 48 (2008).

[32] M. Z. Yao, R. E. Rice, and K. Wallis, *Predicting user concerns about online privacy,* Journal of the American Society for Information Science and Technology **58**, 710 (2007).

[33] E. A. Rose, *An examination of the concern for information privacy in the new zealand regulatory context,* Information & Management **43**, 322 (2006).

[34] G. R. Milne and M. J. Culnan, *Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices,* Journal of Interactive Marketing **18**, 15 (2004).

[35] R. L. Raschke, A. S. Krishen, and P. Kachroo, *Understanding the components of information privacy threats for location based services,* Journal of Information Systems (2014).

[36] C. Van Slyke, J. Shim, R. Johnson, and J. J. Jiang, *Concern for information privacy and online consumer purchasing,* Journal of the Association for Information Systems **7**, 16 (2006).

[37] S. Faja and S. Trimi, *Influence of the web vendor's interventions on privacy-related behaviors in e-commerce,* Communications of the Association for Information Systems **17**, 27 (2006).

[38] C. M. Angst and R. Agarwal, *Adoption of electronic health records in the presence of privacy concerns: the elaboration likelihood model and individual persuasion,* MIS quarterly **33**, 339 (2009).

[39] B. Bulgurcu, *Antecedents and outcomes of information privacy concerns in online social networking: A theoretical perspective,* all sprouts content **10** (2010).

[40] H. J. Smith, T. Dinev, and H. Xu, *Information privacy research: an interdisciplinary review,* MIS quarterly **35**, 989 (2011).

[41] J. Van Den Hoven, *Information technology, privacy, and the protection of personal data,* Information technology and moral philosophy , 301 (2008).

[42] M. J. Culnan, *" how did they get my name?": An exploratory investigation of consumer attitudes toward secondary information use,* Mis quarterly , 341 (1993).

[43] S. Bellman, E. J. Johnson, S. J. Kobrin, and G. L. Lohse, *International differences in information privacy concerns: A global survey of consumers,* The Information Society **20**, 313 (2004).

[44] G. R. Milne, J. Beckman, and M. L. Taubman, *Consumer attitudes toward privacy and direct marketing in argentina,* Journal of Direct Marketing **10**, 22 (1996).

[45] M. L. Katz and C. Shapiro, *Systems competition and network effects,* The Journal of Economic Perspectives , 93 (1994).

[46] A. H. Maslow, *A theory of human motivation.* Psychological review **50**, 370 (1943).

[47] M. Nagenborg, *Designing spheres of informational justice,* Ethics and information technology **11**, 175 (2009).

[48] A. Barth, A. Datta, J. C. Mitchell, and H. Nissenbaum, *Privacy and contextual integrity: Framework and applications,* in *Security and Privacy, 2006 IEEE Symposium on* (IEEE, 2006) pp. 15–pp.

[49] G. J. Nowak and J. Phelps, *Understanding privacy concerns. an assessment of consumers' information-related knowledge and beliefs,* Journal of Direct Marketing **6**, 28 (1992).

[50] P. A. Pavlou, *State of the information privacy literature: where are we now and where should we go,* MIS quarterly **35**, 977 (2011).

[51] F. Bélanger and R. E. Crossler, *Privacy in the digital age: a review of information privacy research in information systems,* MIS quarterly **35**, 1017 (2011).

[52] R. Gross and A. Acquisti, *Information revelation and privacy in online social networks,* in *Proceedings of the 2005 ACM workshop on Privacy in the electronic society* (ACM, 2005) pp. 71–80.

[53] J. Phelps, G. Nowak, and E. Ferrell, *Privacy concerns and consumer willingness to provide personal information,* Journal of Public Policy & Marketing **19**, 27 (2000).

[54] M. Zviran, *User's perspectives on privacy in web-based applications.* Journal of Computer Information Systems **48** (2008).

[55] S. J. Milberg, H. J. Smith, and S. J. Burke, *Information privacy: Corporate management and national regulation,* Organization science **11**, 35 (2000).

[56] W. Hong and J. Y. Thong, *Internet privacy concerns: an integrated conceptualization and four empirical studies,* MIS Quarterly **37**, 275 (2013).

[57] S. H. Rackow, *How the usa patriot act will permit governmental infringement upon the privacy of americans in the name of" intelligence" investigations,* University of Pennsylvania Law Review , 1651 (2002).

[58] A. Westin, *Privacy and Freedom* (Bodley Head, 1970).

[59] Qualtrics, *Understanding your data set,* (2014).

[60] P. A. NORBERG, D. R. HORNE, and D. A. HORNE, *The privacy paradox: Personal information disclosure intentions versus behaviors,* Journal of Consumer Affairs **41**, 100 (2007).

[61] Y. Tsarenko and D. R. Tojib, *Examining customer privacy concerns in dealings with financial institutions,* Journal of Consumer Marketing **26**, 468 (2009).

[62] M. J. Culnan, *Consumer awareness of name removal procedures: implications for direct marketing,* Journal of direct marketing **9**, 10 (1995).

[63] N. F. Awad and M. Krishnan, *The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization,* MIS quarterly , 13 (2006).

[64] D. George and P. Mallery, *Spss for windows steps by steps: A simple guide and reference 11.0 update.* (2003).