Marja Vatka

INFORMATION BEHAVIOUR and DATA SECURITY

Health Belief Model Perspective

Master's Thesis in Governance of Digitalization
Supervisor: Docent Shahrokh Nikou
Faculty of Social Sciences, Business and Economics
Åbo Akademi University

Åbo 2019

ÅBO AKADEMI UNIVERSITY – Faculty of Social Sciences, Business and Economics

Subject: Governance of Digitalization
Writer: Marja Vatka
Title: INFORMATION BEHAVIOUR AND DATA SECURITY – Health Belief Model Perspective
Supervisor: Shahrokh Nikou
Abstract: As the use of different applications and platforms has increased together with the use of Internet-connected devices, concerns regarding user's safety have risen. Allowing the service provider to use the user's location and personal information, the user can have more optimised results based on the given information when searching for information. The collected data can be shared or sold further to third parties by the service provider. This can cause concern for the user, as the user might not know, which services have his personal information and for what purpose the information is used. This concern has been acknowledged and new measurements for protecting one's safety have been established. New legislations have been executed to protect the data that users provide for the used services. According to the European Union's General Data Protection Regulation (GDPR), users have more rights concerning their data. This thesis uses quantitative method to investigate user's behaviour concerning data security by adopting Health Belief Model (HBM) and proposing a conceptual framework based on HBM. The variables in the HBM including Perceived Susceptibility, Perceived Benefits, Perceived Barriers, Cues to Action, Self-efficacy and Perceived Seriousness were used in order to find out the underlying beliefs towards data security behaviour. 133 respondents participated in an online survey. Partial Least Squares Structural Equation Modelling (PLS-SEM) technique is applied to analyse the data. Findings show that differences between age and gender can be detected when it comes to data security behaviour. Perceived Seriousness can be argued to have an effect on Perceived Susceptibility, Perceived Benefits, Cues to Action and Self-efficacy and thus indirectly to Data Security Behaviour.
Date: 16.5. 2019 — Number of pages: 80
Keywords: Information Behaviour, Data Security, Health Belief Model, user perception

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincerest gratitude for my thesis supervisor Shahrokh Nikou, Docent, Information Systems Åbo Akademi University, who supported me in all the stages when writing this thesis. His observations and guidance were vital throughout the research and writing process.

I am thankful to all my beloved friends and respondents of the survey, who provided me with good feedback and advice. I also wish to thank my fellow students, who made the study time in Åbo Akademi University worthwhile. I will look back on fondly the fun times and interesting discussions we shared and hope to keep in touch with you as we continue our own from now on.

Most importantly, I want to thank my family and closest supporters. My mother and father for always believing and encouraging me through thick and thin, especially in my time of need and my brothers for expressing their confidence in me. I hope that the loved ones who have departed this life are proud of my achievement and join me in my celebration from afar.

Marja Vatka
Åbo, Finland
May 2019

ABBREVIATIONS LIST

EU	European Union
GDPR	General Data Protection Regulation
HTML	Hypertext Markup Language
URL	Uniform Resource Locator
LBS	Location-based Services
GPS	Global Positioning System
EEC	European Economic Community
DPO	Data Protection Officer
IR	Information Retrieval
CMIS	Comprehensive Model of Information Seeking
HBM	Health Belief Model
ISA	Information Security Awareness
IS	Information Systems
ISP	Information Security Policy
TAM	Technology Acceptance Model
TRA	Theory of Reasoned Action
BEH	Data Privacy Behaviour
PSS	Perceived susceptibility
PB	Perceived benefits
PBR	Perceived barriers
SE	Self-efficacy
PSE	Perceived seriousness
CA	Cues to action
MGA	Multi-Group Analysis

1 INTRODUCTION

Today's society is mostly based on the usage of the Internet. It is present in whatever we do, and it is extremely difficult to conduct our work, studies or issues concerning our health or finance without it. As the European Union's (EU) General Data Protection Regulation (GDPR) reaches its first year in being active, the user perception concerning rights over their personal security and personal data and the users' knowledge is to be observed. One of the main reasons for the data protection legislation renewal has been to more carefully conduct privacy protection. The principal purpose of privacy protection is to identify private information depending on the presented context and applied law.

Personal data is important for e-business for various reasons. Even though people are more aware of their information being used for other than its original purpose, all information we provide online is useful for marketers and intermediaries. Details such us age, sex, family and economic status give marketers a wide picture of consumers' behaviour and online activities. This enables the use of target marketing and personalized sales and offers. Even though consumers have rights for their personal data, so have marketers. In order for marketers to conduct their actions, they have to be up to date about the latest data protection and privacy laws (Chaffey, 2009, p. 210). One could imagine that laws would be read the way they were meant. Yet, in the business world there is room for different interpretations. Often it is up to the manager to make the decision based on their own evaluation and possible outcomes for the company. Companies could still face risks such as declined reputation and loss in revenue if they are caught of faulty compliance of the law. For companies to execute successful business and specifically, e-commerce, they need to be able to adapt harmony between the benefits for the customer's online experience and thus giving out personal information to the hands of the company (Chaffey, 2009, p. 210). This is also closely linked with the ethical side of business. It can be said that ethical standards are personal or business practices or behaviour which are generally considered acceptable by society (Chaffey, 2009, p. 209). Laws developed based on ethics are designed to control the morality of internet marketing. Often laws concerning e-commerce are not developed enough to follow the development of technology.

As of late, misconducting of a user's data have come into light. According to DLA Piper's cybersecurity report, within the first eight months since applying the GDPR, over 59,000 reported data breaches occurred in Europe (DLA Piper, 2019). Arguably one of the biggest issues has been the Facebook and Cambridge Analytica issue, in which over 87 million Facebook users' data was given to the analytics company Cambridge Analytica (Isaak & Hanna, 2018). Even though there are no clear implications of how it did effect on people's use of social media platforms, the question concerning individual rights has been in the headlines ever since.

1.1 Research objectives

This thesis seeks to explain the importance of data privacy and the understanding of data privacy and security from the perception of the user. The importance of data protection and privacy are observed and how they can affect one's understanding over their own rights. This thesis seeks to answer, how information seeking behaviour and security compliances are linked from the perception of the user and what different attributes effect on the use of different services considering data privacy. Also, in the light of issues such as Facebook-Cambridge Analytica, users' attitudes towards platforms and services will be studied.

1.2 Research questions

Based on the objectives of the thesis, three research questions have been formulated to observe the objectives of the thesis. The research questions will provide a structure for the thesis.

The author has stated following research questions for the study:

1. How does perceived seriousness of data privacy risk affect data privacy behaviour?
2. How does personal demographics affect perceptions concerning data behaviour?
3. Which constructs have a significant effect on data privacy behaviour?

1.3 Structure of the research

Research structure provides a synopsis for the study. Presenting the structure helps the reader to better perceive the arrangement of the research and the justifications to it. This study includes eight chapters.

The introduction chapter lets the reader to familiarize with the subject at hand by explaining the background and problem that the thesis explores. Research problem, research questions, objectives and motivation are proposed in order to simplify the understanding of the subject.

Literature review will be observed in the second chapter of the thesis. This chapter introduces the existing literature that the author has used in her justifications and offers more in-depth knowledge about the subject. Relevant literature concerning the EU's history regarding data security will be discussed, as well as previous studies and user acceptance perception is defined.

In the third chapter, the theoretical background will be presented. The relation between human behaviour and security is presented with research model for investigating human behaviour. Also, other relevant theories are discussed and framework for the study is defined for the thesis

The fourth chapter discusses and validates chosen methods and their application by presenting the conceptual framework of the study. The author will discuss the motivation behind the research. Also, hypotheses for the study are presented.

Research methodology will be presented in the chapter five. The author explains reasoning for chosen research methods and will explain more in detail, how the data for the study is collected and how the data will be analysed.

The sixth chapter explores the analysis of the study. Hypotheses will also be tested, and the author will discuss the results of inner and outer model analysis.

The discussion and results of the study will be presented in chapter seven. Using the framework introduced earlier in the thesis, the results will be discussed more in detail. Thus, chapter six offers more discussion concerning the perceived results.

The final chapter concludes the research. The author will provide suggestions for further studies and possible limitations are observed.

2 LITERATURE REVIEW

This chapter provides relevant background information and theoretical foundation such as security issues, data security legislation and information behaviour.

2.1 General security concerns in the modern era

2.1.1 Physical Security

Almost everywhere you go, one cannot but help to notice that every other passer-by is looking at their mobile device or are somehow linked to one on the go. In the age of mobile devices, people tend to forget how easy it is to steal a mobile phone laying on a cafeteria table or just take it from your hand on the street. The small size and portability increase the risk of losing a device and its content (Walters, 2012). At best, the cost for a lost or stolen device is only money but worst, someone has a quick access to all your personal information. Also, simply borrowing a phone for a call might bring trouble; possible malware or an app is downloaded and installed quickly to a user's phone without them acknowledging it. It could be argued that despite concerns raised by cloud computing, physical security is the most important security risk for mobile devices (Dwivedi et al., 2010).

2.1.2 Phishing and Safe Browsing Environment

Web users are more and more threatened by the severity of phishing. Usually, web users are made to believe that they are using a trustworthy service provider when their account and identity information and logon credentials can be hacked to criminal intentions. Mostly phishing attempts are launched by sending emails with links to fake websites that gather information about the user (Fette et al., 2007; Rogers, 2006). The risk for phishing still exists even on mobile devices. One of the biggest reasons for this is that using a mobile device makes it easier to click on various items on the screen without considering. Hence, numerous mobile Hypertext Markup Language (HTML) browsers are not able to show the Uniform Resource Locator (URL) fully. When the user cannot view the full URL on a mobile browser, strengthens this the possibility of phishing (Dwivedi et al., 2010). It is also difficult for the user to evaluate the trustworthy of a link when it is shared

via SMS, email or social media, in which the user is connected all the time and can just click without examining the origins.

2.1.3 Location Privacy and Security

Services based on location take advantage of positioning technologies in order to offer users access that the users would not be able to have without (Xu et al., 2010). As different services are easily available and many functions can be conducted via mobile, privacy has become harder and harder to point out among users. Privacy is one of those things that is hard to pinpoint with users. There is a paradox as recent news and media have raised the alarm considering security and thus mobile users want to be safe but at the same time donate their personal details by using different services such as sharing their location on Snapchat or other social media service, where your friends can look up one's location, but also give the same location to the service provider to serve their intentions. Location-based Services (LBS) provide more value by taking advantage of user's location information and personal information. Nevertheless, LBS are expected to raise consumer's concern over their privacy as location information could possibly be sensitive for the consumer. (Choi et al., 2007). Sharing location has never been a big of an issue among computer or laptop security. In the mobile context, use of Global Positioning System (GPS) and location sharing raises concerns that have not been around before (Choi et al., 2007; Dwivedi et al., 2010).

2.2 Data security and regulation in the EU

The following sub-chapter describes the general guidelines of data security and its regulation in the EU area. The author brings insight to the process of how an initiative becomes a regulation and how it is adopted, as well as an outlook on the history and basis of data regulation in the EU.

2.2.1 Definition of Data Security

Privacy perception can vary depending on a country, culture or jurisdiction (Chen et al., 2012). Various definitions for data security can be found from literature. Data security can be stated as "protection of data from unauthorized (accidental or intentional)

modification, destruction, or disclosure" (National Institute of Standards and Technology, 2013) and "the measures taken to prevent unauthorized access or use of data." (OECD, 2007) In broad terms, privacy can be said to be linked to collection, use, disclosure, storage, and destruction of personal data (Chen et al., 2012). The fundamental purpose of privacy protection is to identify private information depending on the presented context and applied law. (Chen et al., 2012) Importantly, many of the richest emerging sources of social network data come from settings such as e-mails, instant messages or telephone communication. Users have strong expectations of privacy on such data. For example, in cloud computing, inappropriate security measures regarding data operations and transmissions can set the data at high risk (Rao et al., 2015).

The purpose and use of data can be observed through data life cycle. Data life cycle (Figure 1) covers the process starting from generating data to destruction of the data. Chen et al. (2012) have divided the data life cycle into seven phases.

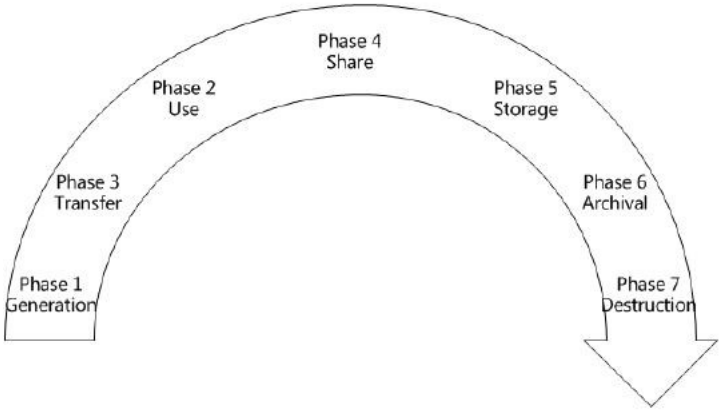

Figure 1. Data life cycle (Adapted from Chen et al, 2012)

From the conventional point of view, users or organizations acting in the field of IT own and manage data. With this ownership, data generation is involved. However, if data should be transferred into cloud, the ownership of the data is not easy to preserve. When it comes to the private information, the owners of the data hold the right to know, what personal information is being collected and also stop collection and use of this private information. (Chen et al., 2012)

When data is transmitted inside organizational limits, complex data encryption measures are seldomly needed. If data is transmitted beyond organizational borders, the confidentiality of the data and integrity need to be secured so that data cannot be altered

by unauthorized users. This being said, data encryption might not be sufficient as data integrity has to be secured as well. Thus, these transport customs need to be support confidentiality and integrity. In regard to transmitting data between different cloud storage services, confidentiality and integrity need yet again be secured. (Chen et al., 2012)

As data encryption might bring problems with indexing and query, static data used by cloud-based applications is not usually encrypted. Also, in conventional IT, the used data is hardly ever encrypted. Processed data in cloud applications is usually stored with other users' data, which might cause threat to unencrypted data. In case of private data, the data owners have to concentrate on the use of their personal information: is it used for the purposes the information was collected and is it shared with external parties. (Chen et al., 2012)

When data is shared, ways to use data are multiple and conducting data permissions comes more complicated. If data owner provides access to his data, this actor can share the data forward to a third party without given consent from the owner of the data. Thus, if data is shared, the third party's interest needs to be considered as it may or may not maintain the authentic protection measures and restrictions for use. One should also keep in mind that the other parties might split the received information depending on the sharing policy. (Chen et al., 2012)

If data is stored in cloud storages, three aspects of information security need to be taken into consideration: confidentiality, integrity and availability. Data encryption is generally used for data confidentiality. Here, the importance of encryption algorithm and key strength are in the centre. In addition to storage and handling, processing speed and computational efficiency of encrypting great deal of data need to be taken into consideration. Users might not always know, where exactly their data is stored. It might be challenging for the users to first download and then upload the data to verify the integrity of data in the cloud. Also, traditional technologies might not be sufficient in detecting data integrity. Data availability might be threatened due to external attacks but also cloud providers future, availability and probability for backup might raise concerns for the user (Chen et al., 2012).

Archiving data and availability of data in cloud might be under threat if the cloud service providers do not offer off-site archiving. Availability and privacy threats might occur if the storage duration is not persistent with archival requirements (Chen et al., 2012).

The final stage of the data life cycle involves destructing the data. Even though the user might have thought the data is deleted, due to physical characteristics of the storage medium the data might still prevail and be restored. Understandably, this might lead to unintentional disclosing of personal, delicate information (Chen et al., 2012).

2.2.2 History of data regulation in the EU

Ever since the European Union was established in 1993, the political and economic union has aimed to bring stability and unity across the Europe. The 28 Member States work together in 35 different policy areas which influence their citizens, such as migration, economy, business and education (European Union, 2018, p. 11) The idea of a coalition between the European countries was founded after the Second World War when the world was getting back to its feet after a long and tumultuous period. The primary rationalization was that if countries practice trade with one another, they are more dependent from other countries and less likely willing to cause conflict. The result was the predecessor of the EU, the European Economic Community (EEC), founded in 1958 (European Union, 2018).

Even though the EU was based on making trade easier between its member countries, the companionship and harmony is a never-ending process, which is to be maintained with united laws and regulations among all the states joined in the EU. As the world changes and societies evolve, so do the regulations set by the EU. Data regulation and security is one of the key issues in today's world, which also the EU has acknowledged. The first directive for data security was set in 1995, after it was perceived that the data protection standards needed to be adjusted to support data transfer within the EU and across borders. Different countries had their own rules, which varied enormously from one another with their level of protection from the legal perspective – both for consumers and data processors (Voigt et al., 2017, p. 2). Subsequently, Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data of the European Parliament and the Council of 24th of October 1995 was

endorsed. The directive regulated, how one's personal data is to be processed and ensured the free flow of this data between the EU Member States. Directives do not act as laws directly and thus need to be converted to national laws by each Member State. As this was not executed successfully, the Data Protection Directive deteriorated and legal differences emerged among Member States; the directive was interpreted one way in one country and another way in another country, which caused different unlawful activities in different countries (Voigt et al., 2017, p. 2).

It was not until 17 years later in 2012 when the initial proposal for updating data protection regulation was made by the European Commission. By 2015 the Parliament and Council reached an agreement after various suggestions and discussions to improve the data processing and data security within the EU.

2.2.3 General Data Protection Regulation

In the European Union, the greatest change in data privacy in over 20 years occurred when the new General Data Protection Regulation (GDPR) came into effect in May 2018. In the European Union directive from 1995, directive 95/46/EC states that the directive "sets up a regulatory framework which seeks to strike a balance between a high level of protection for the privacy of individuals and the free movement of personal data" (EURlex, 2018). Understandably, some of the issues regulated by this directive are outdated, which is why new regulation was put in action starting from May 25 2018 (EUGDPR, 2018). The regulation has brought changes to the world of business since the regulation is applied to all companies which process the personal data of data subjects living in EU, no matter where the company is located. (EUGDPR, 2018) The following lists the data subject rights set by GDPR.

Breach Notification

Breach notification is compulsory in all EU member states where "a data breach is likely to result in a risk for the rights and freedoms of individuals". The notification has to be done within 72 hours of occurred breach. Also, data processors role is accentuated as they are obligated to notify customers as soon as is practicable after becoming aware of a data breach (EUGDPR, 2018).

Right to access

As one of the points in the GDPR, the data controller needs to be more transparent regarding the obtained information they have of the data subject. The data controller needs to provide the data subject the information, whether or not the data subject's personal data is being processed and if so, where and for what purpose. When asked, this information needs to be provided free of charge and in an electronic format (EUGDPR, 2018).

Right to be forgotten

The right to be forgotten, also known as Data Erasure, ensures that the data subject can make the data controller remove their personal data. This also covers disabling third parties' use of the data and further sharing of the data. From the data controllers' point of view this means, that they need to erase data that is no longer relevant for the original purpose it was collected. Also, if the data subject withdraws their consent, the information needs to be erased (EUGDPR, 2018).

Data Portability

According to data portability, the data subject has the right to receive all personal data related to them and have the right to transfer the data to another data controller (EUGDPR, 2018).

Privacy by Design

As a concept, privacy by design is not anything new. Nevertheless, it was not until the data security renewal that it was taken as a part of a legal requirement. The main point in privacy by design is that the data protection aspect would be considered as a crucial part right when a system is designed. Article 23 related to privacy by design also states that data controllers should hold and process data that is unquestionably important for executing their activities and also restrict the access to personal data for those who are the actual processors (EUGDPR, 2018).

Data Protection Officers

Set by the GDPR, controllers and processors whose main activities are processing operations, where regular and systematic monitoring of data subjects on a large scale or special categories of data or data relating to criminal conviction need to appoint a data

protection officer (DPO) internally. The characterization of a DPO is described as follows:

"Must be appointed on the basis of professional qualities and, in particular, expert knowledge on data protection law and practices; May be a staff member or an external service provider; Contact details must be provided to the relevant DPA; Must be provided with appropriate resources to carry out their tasks and maintain their expert knowledge; Must report directly to the highest level of management; Must not carry out any other tasks that could results in a conflict of interest" (EUGDPR, 2018).

2.3 Information behaviour

As in all aspects in life, there are as many ways of doing things as there are people. Interaction between people and human behaviour is intriguing and thus also a compelling field of study. However, Information Behaviour and Information Seeking were acknowledged as scientific interest mere 50-60 years ago. One could have expected this to be studied even earlier, as there might be different outcomes to the study depending on the underlying research process. Information behaviour has been studied in various contexts and different factors' influence on it have been distinguished. These factors have been categorized as cognitive, ecological, psychological, social, spatial and systemic.

Cognitive approach to information behaviour was applied to to increase the effectiveness of information retrieval process. Afzal (2011) says that two cognitive approaches can be distinguished; user-centred and socio-cognitive (Afzal et al., 2011). The user-centred approach is purely cognitive and focuses on mental models. These models are important when designing Information Retrieval (IR) systems. The focus here is that if the user's mental model can be understood, the better IR systems can be designed. Yet, whether an information system is successful or not, the user's needs have to be presented precisely (Afzal et al., 2011). The socio-cognitive approach's basis was founded as researches argued that the social context of IR was not taken into consideration. Knowledge structures vary between people due to person's take on the surrounding world. These knowledge structures are developed when people are interacting in different environments i.e. cultural, political, social and economic (Afzal et al., 2011).

Heinström's study (2006) observes students and incidental information acquisition regarding personality, as well as study approaches. The author mentions that previous studies had issued that feelings of confidence, relief, optimism and satisfaction are linked with increased incidental information acquisition. Previous studies has also shown that feelings of disappointment, frustration, confusion, uncertainty, and anxiety would in contrast diminish these positive feelings. Heinström (2006) also mentions that key factors here are the importance of used medium as well as person's motivation and curiosity. The factors explain how someone could be more effective than other person and why it is important to support work motivation and being goal-oriented in order to be successful and enjoy.

2007 research conducted by Tötterman and Widén-Wulff (2007) discusses different social aspects and their effect on information behaviour. The research addressed social aspects in university environment, where people co-operate with others from different cultures and nationalities. Their study illustrates the importance of understanding the reasons behind varying opinions, as people's experiences effect on their information behaviour and thus might collide, for example when working with other people.

The principal of proposing spatial approach to information studies and information behaviour is a quite new method. Since the 1990s, user-centred approach gained more interest and thus contextual and situational factors in information seeking became more prominent. Although, Savolainen (2006) mentions that the interest towards spatial factors started already in the 1960s. Along with growing interest, recent networks brought new possibilities for information seeking. Naturally this affected on the way people search for information as people were no longer dependent on a certain location. Further, Savolainen says that spatial factors may also refer to the ways information seeker acknowledges useful information sources. Due to false advertising and communication present in the modern era, this fact cannot be addressed too much.

Johnson and Case (2012) examine elements regarding Comprehensive Model of Information Seeking (CMIS) (see Figure 2), established by Johnson et al. in 1995, and major attributes affecting it. Two models that they have presented to be relevant to the CMIS model are Health Belief Model (HBM) and Transtheoretical Model. The Health Belief Model, founded in the 1950's by the United States Public Health Service's psychologists, (Rosenstock, 1974) aims to explain and predict preventive health

behaviours, while Transtheoretical describes how intentional change of an individual influences on the way information is sought. Transtheoretical Model was developed in the late 1970s by Prochaska et al. The basis of Transtheoretical model lie highly on psychotherapy and behaviour change and aimed to bring together the field, which obtained various theories (Prochaska, 1992).

According to HBM, people try to prevent possible threats to their health as soon as they think they are susceptible to a disease. In CMIS, antecedents affecting information seeking can be divided into four: demographics, personal experience, salience and beliefs. Different beliefs may also act as barriers towards information seeking. Johnson and Case (2012) state that "an Individual's belief in the efficacy of various medical procedures can also affect health information seeking and preventive behaviour" (Johnson and Case, 2012, p. 58). However, empirical tests have not widely distinguished beliefs such as health motivation, locus of control and self-efficacy.

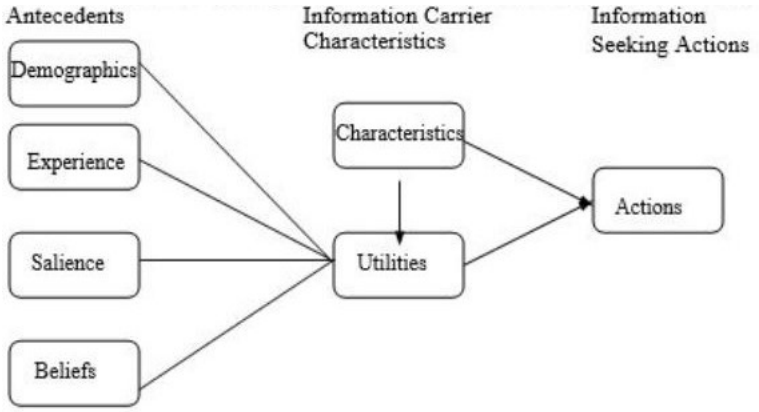

Figure 2. Comprehensive Model of Information Seeking Johnson, J.D. et al, 1995

2.4 Behavioural information security

Behavioural information security can be defined to be focused "on the behaviours of individuals which relate to protecting information and information systems assets which includes computer hardware, networking infrastructure, and organizational information" (Crossler et al., 2012). According to Crossler et al. (2012) the field of Information

Security (InfoSec) research covers various of fields and concerns in protecting and diminishing threats to the information assets and technical resources available within computer-based systems (Crossler et al., 2012). As of late, most of the research has focused on the technical aspect of security and not as much to the socio-philosophical aspect. The research is more and more focusing on the behaviour of individuals and their decision-making (Komatsu et al., 2013). It has been realized that characteristics such as openness to experience, honesty and agreeableness are in correlation with lower risk taking and higher information security awareness (Hadlington, 2017). It could be added that also cultural dimensions have an effect on one's behaviour. For example, 2007 research by Tötterman and Widén-Wulff focused on different social aspects and their effect on information behaviour. One of their findings was that even though each culture has their own social capital, there are also different ways to share information in the community (Tötterman et al., 2007).

2.5 Information security awareness (ISA)

Information security awareness (ISA) has been widely studied in organizational context. In an organizational setting, ISA is defined "as an employee's general knowledge about information, general knowledge about information security and his cognizance of the ISP of his organization" (Bulgurcu et al., 2010, p. 532). As companies are relying more on their information systems (IS), companies and organizations must pay more attention to managing the risks that come along with IS, such as how well their employees are aware of their information security policy (ISP) and what could be possible consequences if these are not to be followed. These risks set challenges to modern organizations (Bulgurcu et al., 2010). Organizations now have recognized that information security is related to human factors as well, in addition to technical problems (Hassel et al., 2004). For example, multi-user computing environment, where various databases and other tools are shared with other users and increased use of personal computers cause more concern to organizations (Thomson et al., 1998). Also, even though employees are often considered to be the weakest link in an organization's information security, they could also be valued as the most important asset when it comes to diminishing risks concerning information security. Companies take drastic measures in order to protect their information and technology resources but also have appraised employees trustworthy. Study conducted by Bulgurcu et al. (2010) shows that conducting ISP and ISA policies effects on the

employees' attitudes and thus proves them to have positive intention towards organization's security (Bulgurcu et al., 2010).

Technological turmoil has brought along serious concerns, that various researches in the field of ISA aim to explain and assure, that more attention needs to be paid to these issues. In general, failures in IT means consequences in various fields of business; loss in sales, lowered customer satisfaction, endangered confidentiality, losing credibility in the eyes of stakeholders and jeopardized workplaces also among the board. When these problems continue, it has impact on company's return and debts. The impact can be seen for months after the actual failure has occurred (Benaroch et al., 2017).

3 THEORETICAL BACKGROUND

For the longest time, the core of information systems research has inevitably been the use and adoption of IT, especially in an organizational and workplace setting. Although remarkable development has been made concerning hardware and software systems, users still do not take full advantage out of the various sophisticated tools available. As Davis et al stated: "Computer systems cannot improve organizational performance if they are not used" (Davis et al., 1989).

The following introduces some theories that have been widely used when researching or describing the way people relate to technological change and which factors in their background or surroundings influence their attitude towards technology.

3.1 Technology Acceptance Model (TAM)

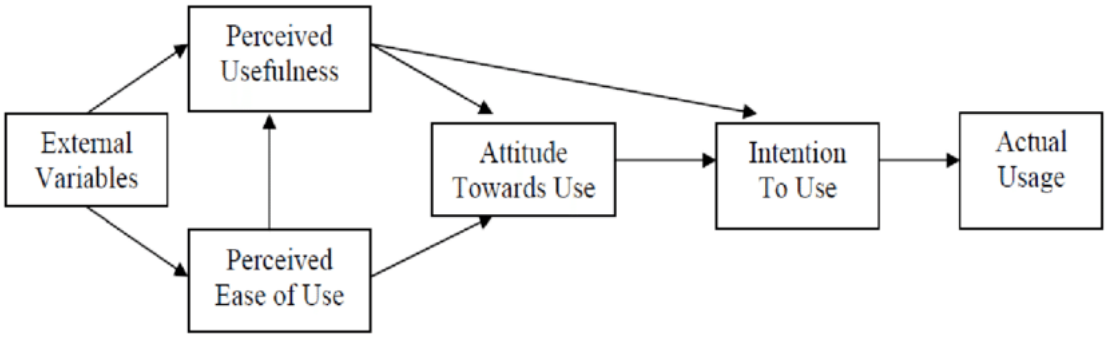

Figure 3. Technology Acceptance Model (TAM), Adapted from Davis 1985

Numerous studies in the IS research area have focused on exploring attitudes and behavioural beliefs. Probably one of the most used models for explaining this is the Technology Acceptance Model (TAM) (see Figure 3), developed by Davis in 1985.

In the 1980s, when Davis developed his model, computers were already the main tool in businesses almost all around the world. Davis started to develop a model, that would aim for two issues. Firstly, propose a model that would reform the understanding of user acceptance process by providing new theoretical observations to how to successfully design and implement an IS. Secondly, Davis hoped to provide theoretical basis for a practical "user acceptance testing" so that system designers could guesstimate new systems before they are implemented and gain useful information considering the users'

perspective (Davis, 1985). All in all, the model he planned was intended to represent "the motivational processes that mediate between system characteristics and user behaviour" (Davis, 1985, p. 10). The key thought in the model was that the overall attitude of a user towards using a system is in relation to whether the user is going to use a certain system in his or her work. Consecutively, attitude consists of two subjects: perceived usefulness and perceived ease of use.

Perceived usefulness, according to Davis can be defined as "the degree to which an individual believes that using a particular system would enhance his or her job performance" (Davis, 1985, p. 26). Perceived ease of use is then again defined as "the degree to which an individual believes that using a particular system would be free of physical and mental effort" (Davis, 1985, p. 26). Davis argues that perceived ease of use has a direct response to perceived usefulness. This is because, according to Davis, more user-friendly system will conclude as increase in job performance. If even a small part of a user's job description includes using a system and if he or she becomes more productive in using that system, the overall productivity will increase. Hence, components of a system could ambiguously have an influence on the usefulness by affecting ease of use (Davis, 1985).

Even though TAM is one of the most fundamental and widely used models in explaining user behaviour and technology adoption, it has some major limitations in fully explaining and capturing individual's behavioural intention. Although perceived usefulness and perceived ease of use are the factors that help to observe acceptance and use of various IT, the beliefs might not fully explain users' relationship towards newly emerging IT, i.e. Internet banking (Wang et al., 2003).

3.2 Theory of reasoned action and theory of planned behaviour

Theory of planned behaviour is largely based on the theory of reasoned action, which was first introduced in 1975 by Fishbein and Ajzen. Fishbein and Ajzen also had a huge impact on previously introduced Davis' research (Davis, 1985, p. 13).

Fishbein and Ajzen have been working together and separately among behaviour prediction and change of behaviour for over 45 years. Over the years, the theoretical

blocks have been altered and in 1975 they stated that the underlying beliefs determine attitudes. When their second book was published in 1980, the model was given a name, a Theory of Reasoned Action (TRA) (see Figure 4). By 1980, the two researchers had developed a standard of producers, which could evoke notable behavioral and normative beliefs and to measure them accoding to their theory. The theory also took into account background factors like demographic, personality and various individual variables. Thus they presented that numerous characteristics in our background lead to shaping behaviour indirectly by shaping the behavioural and normative beliefs the person obtains (Fishbein et al., 2010).

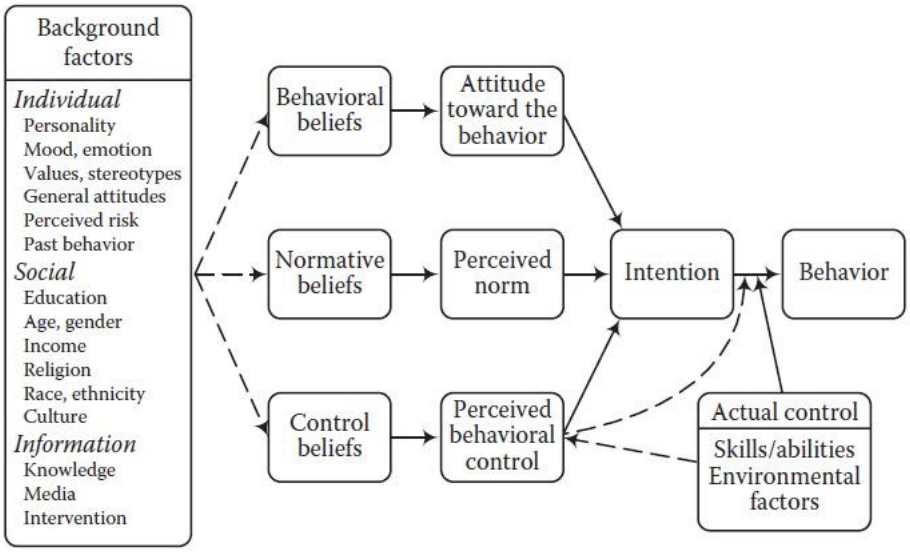

Figure 4. The Theory of Reasoned Action model (TRA), Adapted from Fishbein & Ajzen (1980)

Based on previous collaboration of Fishbein and Ajzen, Ajzen extended the theory of reasoned action. In Theory of Planned Behaviour (see Figure 5), Ajzen aimed to improve the preceding model considering the limitations such as attitudes versus norms and when intention to act is performed, a person can act without limitations. In the centre of the theory is still the individual's intention to perform certain behaviour. Ajzen states, that intentions "are indications of how hard people are willing to try, of how much of an effort they are planning to exert, in order to perform the behaviour" (Ajzen, 1991, 182). In this fashion, when the intention to execute a behaviour is strong, it is expected that the intention is performed. Nonetheless, the behaviour needs to be intended, meaning that he can choose whether or not to operate the behaviour.

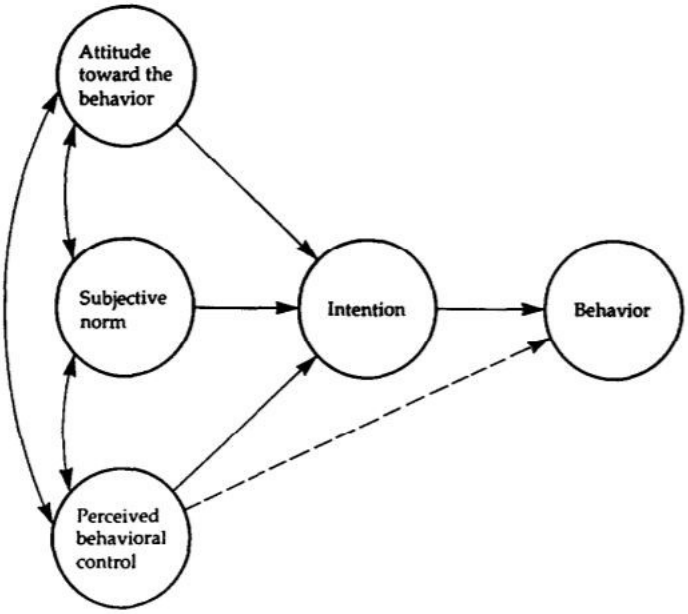

Figure 5. Theory of Planned Behaviour (Ajzen, 1991)

Factors such as time, money or skills can influence the behaviour. The performance of behaviour is highly linked to the opportunities and resources available but also the intention of operating based on a behaviour needs to be active in order to be successful in performing a behaviour (Ajzen, 1991). In their study, Heirman et al. (2013) presented a framework explaining different ways adolescents deal with disclosing their personal information online. Their framework was widely based on the theory of planned behaviour and showed how important social factors are in this setting.

3.3 Research model for investigating human behaviour related to computer security

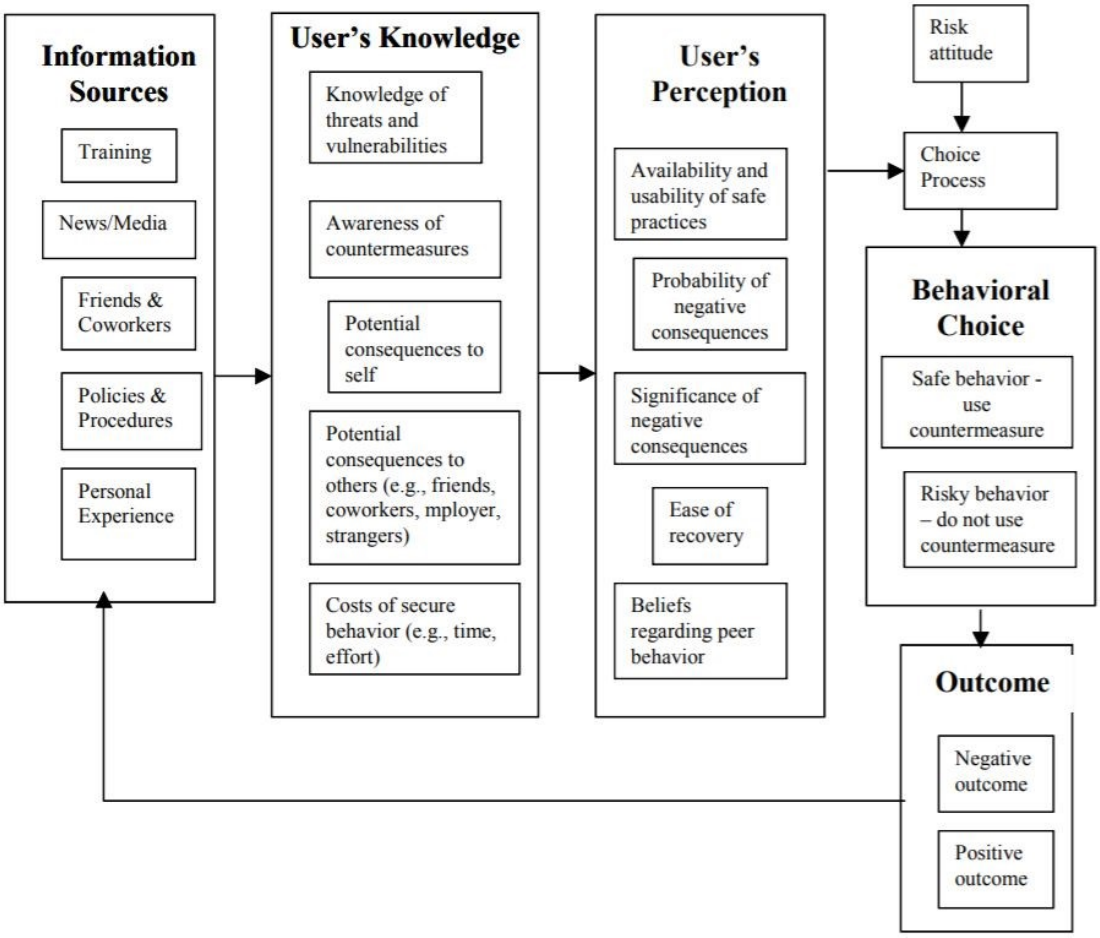

Figure 6. Research model for investigating human behaviour related to computer security (Aytes et al., 2003)

The model for investigating human behaviour related to computer security by Aytes and Conolly (2003) (see Figure 6) again repeats the same, formerly studied important factors affecting performed behaviour. In this model, different factors affecting user's perception such as tendentious media and former personal involvement and knowledge are considered to be important when it comes to the choice process in securing computer activity. The authors assume that the user's action will end up in either safe practice (certainty of no negative repercussions but with costs such as time or effort) or risky practice (no extra costs but with the probability of negative repercussions) (Aytes et al., 2004).

The model especially weighs in following factors: Availability and usability of safe practices. In addition to being aware of countermeasures, they need to acknowledge that they are easy to retrieve. Probability of negative consequences; computer users need to recognize potential threats will be realized. Significance of negative consequences; if a threat occurs, the user will face momentous consequences to themselves and others. Costs of secure behaviour, costs in relation to gained benefits regarding computer security applies both the individual and also organization i.e. cost in time to scan for viruses. Beliefs about peers' behaviour, one is more likely to change their behaviour according to the other people's activity (Aytes et al., 2003).

3.4 Health Belief Model

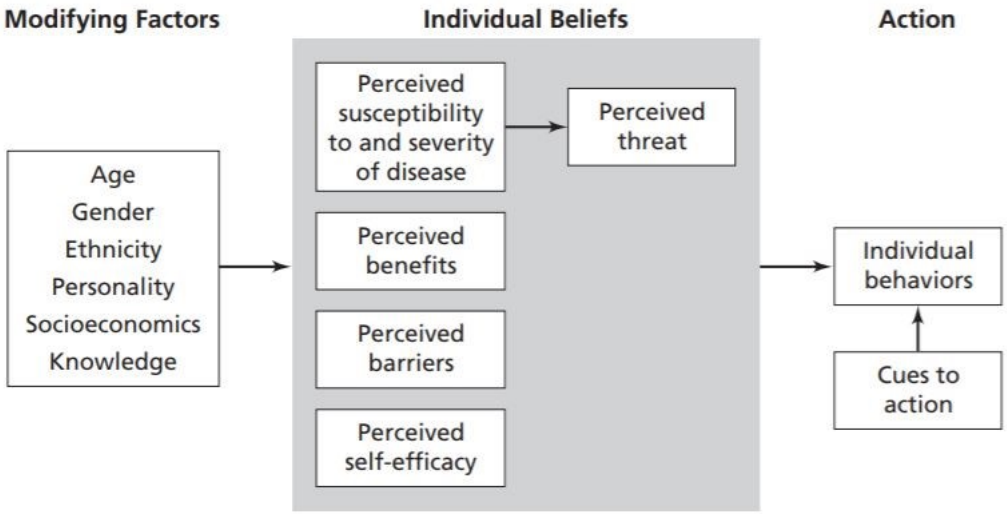

Figure 7. The Health Belief Model Adapted from Becker et al., 1974 (Source: Glanz et al., 2010, 49)

In the 1950s the United States Public Health Service paid much attention to preventing a disease and not much for curing it. At the time, patient's symptoms, aftercare or doctor-patient communication was not valued as high in the Public Health Services as nowadays. Even though preventive measures for such illnesses as cervical cancer, influenza or dental diseases were available, people did not seem to find their way to these preventive tests. Soon it was realized that more research needs to be performed in order to spread public knowledge of the importance of these methods. This led to a group of psychologists in the Public Health Service to conduct a research and form a model, relying highly on

cognitive theories, the Health Belief Model (see Figure 7) while solving practical problems in healthcare (Rosenstock, 1974).

The first traits of the HBM were that for one to act and counter an illness, he would have to rest assure that 1) he was susceptible to it and 2) if been afflicted with a disease, there would be at least up to some extent consequences in his life and 3) taking some particular countermeasures would be beneficial and reduce susceptibility to a condition or weaken the impacts so that he would not have to face barriers such as cost, convenience, pain or embarrassment (Rosenstock, 1974).

The following describes the constructs used in the model, that will predict the reason for people action to prevent, to screen for, or to control illness conditions. (Glanz et al., 2008, p. 47).

Perceived Susceptibility: Perceived susceptibility refers to beliefs or subjective risks of developing a disease. A good example for this is that a woman should be in the belief of having the possibility of getting breast cancer before she is willing to attend a mammogram (Glanz et al., 2008; Rosenstock, 1974).

Perceived Seriousness: As with every adversity, people tend to relate to them differently. The seriousness of an illness could be measured by the emotional point of view (thought of an illness) and the dilemmas rising along the illness. The consequences can be divided as medical and clinical consequences and social consequences. Questions such as is the disease life threatening or does the illness affect so that one must be away from his work for a couple of days and thus receive less salary might arise and vary with their severity. The consolidation of susceptibility and severity has been named as perceived threat. (Glanz et al., 2008; Rosenstock, 1974).

Perceived Benefits: Although one experiences perceived threat, it will lead to change in behaviour depending on the person's beliefs about perceived benefits of the various available actions for reducing the disease threat (Glanz et al., 2010, p48). Glanz et al. also mention that person might experience other non-health-related perceptions i.e. saving money by quitting smoking; this could have an effect on one's behavioural decisions. So, according to Glanz et al., "individuals exhibiting optimal beliefs in susceptibility and severity are not expected to accept any recommended health action unless they also

perceive the action as potentially beneficial by reducing the threat" (Glanz et al., 2010, p. 47).

Perceived Barriers: In this context, barriers are defined as potential negative aspects of a certain health action. These might act as restrictions for acting according to recommended behaviours. A person might carry out unconscious cost-benefit calculation, so that he will evaluate the expected benefits and perceived barriers of an action. For example, if a treatment is expensive, might cause unpleasant side effects or be inconvenient, cost-benefit analysis is conducted (Glanz et al., 2008; Rosenstock, 1974).

Cues to Action: Early on when the HBM was developed, the importance of some trigger or cue to action to rationalize an action was acknowledged. These could include for example bodily events, reminders from healthcare or news provided by the media. Profound, systematic study of cues to action is lacking as the cue might be as vague as a sneeze or the mere conscious perception of a poster. Cues to action is a relatively new concept in the IS research area and have not been studied in many behavioural researches (Glanz et al., 2010; Ng et al., 2009).

Self-Efficacy: Self-efficacy was not originally part of the HBM. The original model was developed to mainly to study preventive health actions (i.e. mammogram) and was not made for observing more complex behaviours. As the HBM could not clearly distinguish difference between individual behaviours, self-efficacy was added only in 1988, when a group of researchers led by Rosenstock added it as a separate construct, along with original concepts. A person should feel capable enough to overcome barriers that are preventing her from taking action, which in this sense is called being self-efficient (Glanz al., 2010). Alas, self-efficacy influences on choosing activities and settings and if success is to be expected, it may have an effect on coping efforts as well (Bandura, 1977).

Other variables: Such as in previous theories, also a person's other characteristics such as demographic and sociopsychological characteristics might have an effect on perceptions. This may have an influence on health-related behaviour. Issues such as education can be seen to shape behaviour by influencing the perception of HBM constructs (Glanz et al., 2010).

4 CONCEPTUAL FRAMEWORK

This chapter describes the selected model to support the research questions. The author will explain reasonings for selected model and background for selected theories, which were introduced in Chapter 3. Also, hypotheses for the research are presented in this chapter.

4.1 Motivation for chosen theories

The whole concept of data security and securing one's personal information on the internet is relatively new. Studies considering organizational computer security can be found, e.g. from Bulgurcu et al. (2010) and Aytes et al. (2004) and more technical aspects of the security but not so many studies, that would be up to date and could relate to the modern digital era and the behaviour we have obtained now that everything is in our reach, just one click away.

As presented in Chapter 3, the field of IS has many theories and models explaining the way technology is adopted by users. Different security measures such as virus software have been widely studied from the computer security perspective. These security measures regarding user's intention to use security appliances can be implemented by using TAM or theory of planned behaviour (Ng et al., 2009). This being said, present-day research concerning security behaviour has shown disputes in positive technologies and protective technologies (Dinev et al., 2007). This, together with newly raised concern concerning personal data security leaves room to other theories that take a closer look on how these technologies are used.

The health belief model has been proven to be useful when finding out underlying behaviour when it comes to security issues such as safe email use and digital threats (Dodel et al., 2016; LaRose et al., 2008; Ng et al., 2009). Based on these previous studies and their outcomes, health belief model will be applied to conduct a survey supporting hypotheses in this thesis.

4.2 Conceptual model and hypotheses

This thesis study will modify the health benefit model and its use by Ng et al. (2009) in computer security context. In their study, Ng et al. (2009) investigated what attributes effect the user to practice computer security, as it was considered to be important from organizational perspective that individual's security behaviour is according to organization's beliefs. The research showed, how HBM can be applied and how it can detect various underlying perspectives that influence on the user. As the study is 10 years old and the context is over-studied, the author aims to repeat the study but using different constructs. In this thesis, the constructs in HBM, Perceived Susceptibility, Perceived Seriousness, Perceived Benefits, Perceived Barriers, Cues to Action and Self-Efficacy will be used since according to studies presented before the constructs within the HBM have been proven to be relevant regarding behaviour and preventive measures (Jayanti et al., 1998). As the main focus of the study is how a possible threat effects on individual beliefs. Perceived Seriousness will be used as an antecedent variable to constructs of HBM which has also a direct relation to security behaviour. The research model is illustrated below (see Figure 8).

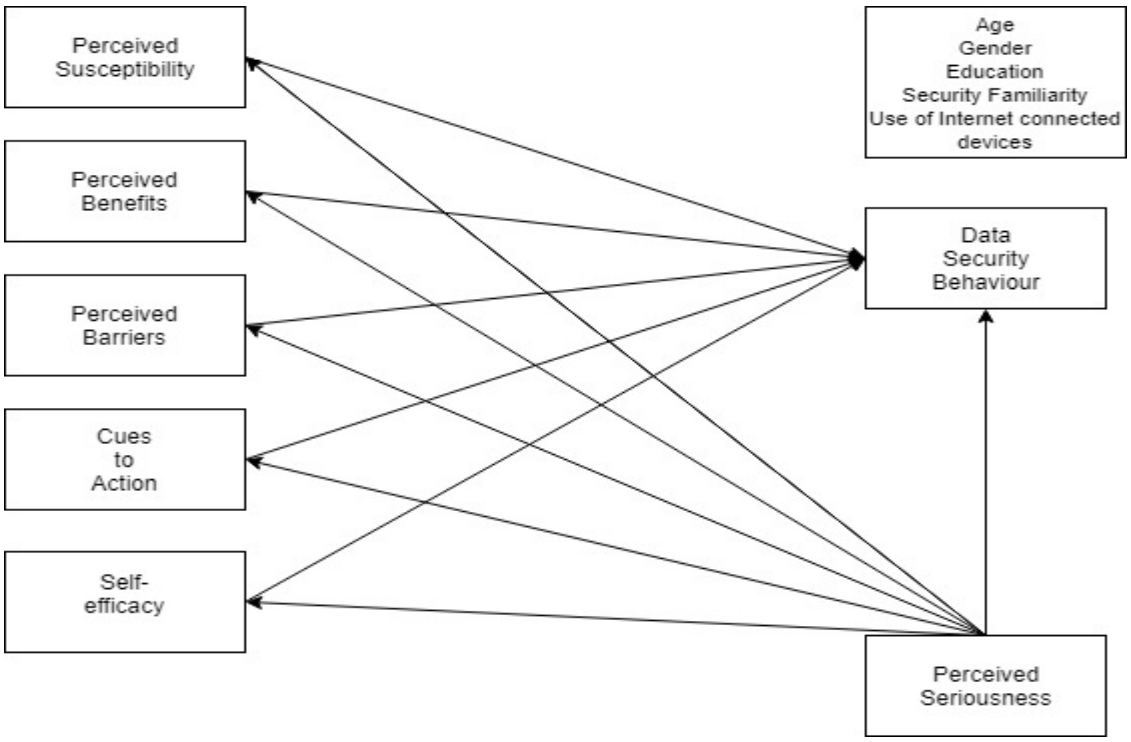

Figure 8. Proposed research model

4.2.1 Perceived Susceptibility

Ng et al. (2009) concluded that Perceived Susceptibility, Perceived Benefits, and Self-Efficacy were important factors in individuals' computer security behaviour (Ng et al., 2009). As stated in the theory section, Perceived Susceptibility covers the beliefs or subjective risks of developing a disease. On one hand, a person might be more cautious and thus feel threatened by an illness when on the other hand one might deny the possibility of falling ill – even though both are offered the same information and facts about the sickness. The author argues that the same complies with people concerning their own data security. If one sees preeminent susceptibility to security incidents, one is more likely to perform more counter measures according to his data security behaviour. Therefore, the following hypothesis is proposed in this thesis.

H1. Perceived susceptibility to security incidents is positively related to data security behaviour.

4.2.2 Perceived Benefits

Person's behaviour will change if there will be some Perceived Benefits in adapting new behaviour, such as in the example about giving up smoking and saving money. Also, person's social groups might result in beliefs considering the gained benefits (Rosenstock, 1974). In this context, perceived benefits indicate to perceived security by practicing data security measures. The author hypothesizes:

H2. Perceived benefits of practicing data security measures are positively related to data security behaviour.

4.2.3 Perceived Barriers

Perceived Barriers could act as restrictions considering person acting according to recommended behaviour. Even though person might feel that a certain action is powerful in reducing threat, the action in question might cause him unnecessary pain or other inconvenience. This could be issues such as two phased-authentications to log in to Facebook or securing your online shopping made with credit card with your online banking codes. Also, costs from installing a virus software could be counted as a barrier. These barriers are likely to reduce performing data security behaviour. The following hypothesis is proposed in this thesis:

H3: Perceived barriers of practicing data security measures are negatively related to data security behaviour.

4.2.4 Cues to Action

In order for a person to start acting and securing his security or preserve his health, a Cue to Action has been proven to be needed. Also, if one has previously been afflicted, the user might detect upcoming concerns easier (Dodel et al., 2016). Here the author will observe, how news, media, social circles or earlier experiences might act as a trigger. The following hypothesis is proposed in the study:

H4: Cues to action are positively related to data security behaviour.

4.2.5 Self-Efficacy

Self-Efficacy states that a person should have the capability and confidence to overcome barriers that are disabling him from taking a particular action. It is related to one's capability and trials for changing his unhealthy or unbeneficial behaviour to better. The author argues that with self-efficacy, one can indeed improve his behaviour towards more successful data security measures. The following hypothesis is proposed in this thesis:

H5: Self-efficacy is positively related to data security behaviour.

4.2.6 Perceived Seriousness

In HBM context, Perceived Seriousness or Severity is related to a person's outlook over the severity of a health issue. Not only does it mean actual harm to one's health, it also covers possible effects on one's financial issues and social relations. In computer security, perceived severity has been stated to be linked to perceived seriousness of a security incident and the consequences it might bring to one's work or the organization. When considering the individual's data security, it could be argued that perceived threat for data security affects the security behaviour in a positive way. The following hypothesis is proposed in this thesis:

H6: Perceived seriousness of security incidents is positively related to data security behaviour.

As the author's original intention was to study the relation of possible threats to our data security and the behaviour affecting people's actions towards it, it is reasonable to observe the Perceived Severity's impact on the other variables. It could be argued that the main goal in practicing safe data security behaviour is to avoid unfavourable outcomes. Thus, the author hypothesizes the following:

H6a. Perceived seriousness of security incidents increases the positive effect of perceived susceptibility on data security behaviour.

In the theory section it was stated that perceived seriousness has an effect on the perceived benefits and perceived barriers. In this context, the author states that once the possible severity is detected, it reduces the user's perception over the threat's effects of perceived benefits and perceived barriers. The statement also relies on the assumption, that if a negative outcome is thought to be serious, the user will conduct safety measures though he might not think that these measures really are that affective and that some kind of protection is needed. In a very serious case, the outcome will also influence so that the user does not mind the costs if he is protected (Ng et al., 2009). The following hypotheses are proposed in this thesis:

H6b. Perceived seriousness of security incidents reduces the positive effect of perceived benefits on data security behaviour.
H6c. Perceived seriousness of security incidents reduces the negative effect of perceived barriers on data security behaviour.

Regarding earlier notes, cues to action are more likely to kickstart an action in a person once a threat is perceived. Thus, it seems that perceived seriousness and cues to action obtain a linkage, which then again affects user's behaviour. So, the following hypothesis is proposed:

H6d. Perceived seriousness of security incidents increases the positive effect of cues to action on data security behaviour.

Lastly, it is argued that perceived seriousness reduces the effect of self-efficacy in this model. No matter the skills the user has, he is still willing to conduct safety measures if

the perceived threat is believed to be harmful, despite he might not trust his capability to do so. The following hypotheses is proposed in this thesis:

H6e. Perceived seriousness of security incidents reduces the positive effect of self-efficacy on data security behaviour.

4.3 Integration of hypotheses

Figure 9 illustrates how the hypotheses are integrated with the proposed research model. The first six hypotheses observe the connection between standard HBM constructs and data security behaviour. The latter five hypotheses shows, whether the perceived seriousness affects negatively or positively in relation to the other hypotheses. Demographic variables such as gender, age, education, internet use and familiarity with security practices will be used to test differencies within each groups.

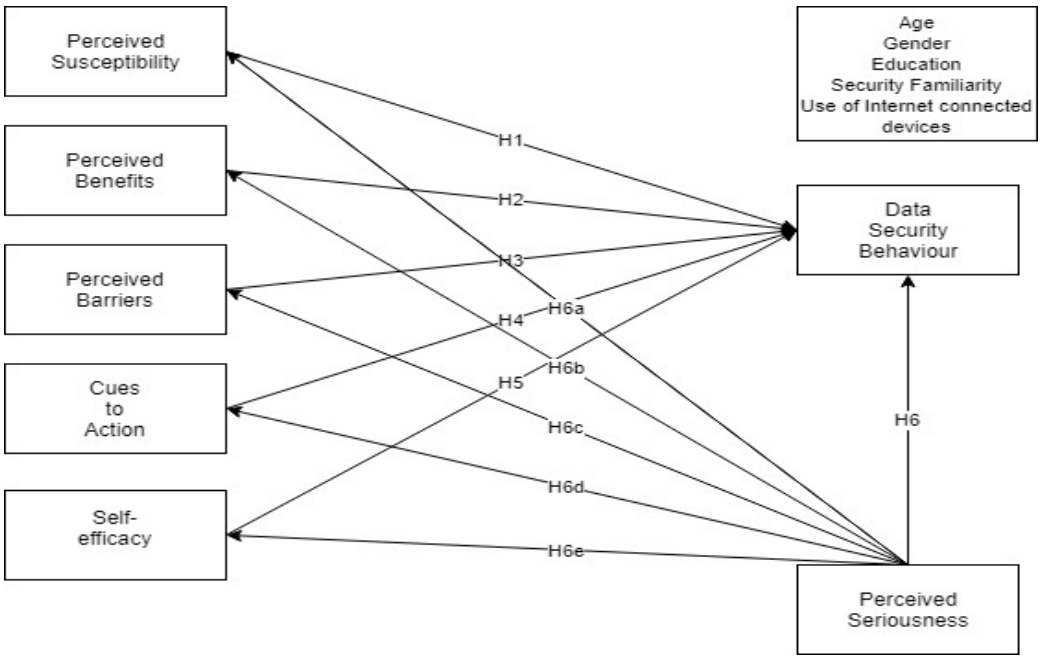

Figure 9. The integration of formulated hypotheses with the proposed research model

5 RESEARCH METHODOLOGY

As discussed above, relevant literature and academic articles have been studied and based on these observations, a conceptual framework has been developed. Testing the formulated hypotheses in the previous chapter in an empirical study will be executed. Examining the relation between different variables is central when answering proposed questions and hypotheses using surveys (Creswell, 2014). As the user perception is in the centre, data will be gathered with a questionnaire and thus the author shall conduct a survey concerning the public's opinion and understanding concerning their data privacy. The questionnaire will be conducted via electronic platform. In order to interpret the answers and analyse them, appropriate software SmartPLS will be used, so that the relations between answers and respondents can be thoroughly observed and hypotheses can be examined. The software SmartPLS was first launched in 2005 and has gained wide popularity thanks to its user-friendly approach and various reporting possibilities. (Wong, 2013)

5.1 Research Method

When conducting a research, the researcher needs to follow a plan or a design. Research methods often go hand in hand with different kinds of research designs. Research design provides a structure which leads the use of a research method and analysis of data. Indeed, a research method is needed to start data collection. The chosen research method will determine the way data is collected and analysed (Bryman, 2012, p. 45).

5.1.1 Quantitative method

Quantitative research can broadly be defined "as entailing the collection of numerical data, as exhibiting a view of the relationship between theory and research as deductive" (Bryman, 2012, p.160). Quantitative research purposes can be divided into three categories: exploratory, descriptive and explanatory. Exploratory clarifies a problem and makes easier to understand a certain phenomenon. Descriptive aims to interpret i.e. a certain event, whereas explanatory explains relations between variables.

5.1.2 Qualitative method

According to Dawson (2002), qualitative research "explores attitudes, behaviour and experiences" (Dawson, 2002, p. 22). Qualitative method also aims to discover concealed motives and ambitions of the respondent, discovering the underlying motives and desires (Kothari, 2004). In order to study these, survey is used as a source for the research. As the qualitative method observes respondents' personal attributes, the number of participants might be low but even more beneficial.

5.2 Methodological choices of the research

It can be said that all research takes advantage from preceding researches and the knowledge they have created. The author has provided reasoning for selected research method by introducing the ground founding theories in the field of information behaviour and how they have been developed further. The literature review has also provided background for the study, as concepts and context were introduced.

The empirical part of the thesis concerns the study, which tries to explain the importance of different personal attributes set by HBM and test how HBM works in the presented context. Quantitative research methods are used to explain how the different attributes are affected by perceived seriousness of a security threat and how all the attributes effect on data security behaviour.

5.3 Data collection method

The survey will be conducted through a web-based survey platform. Participants of the survey are hoped to answer to the questions truthfully, so that it would provide the study with credibility. As a web-based survey is easy to share and answer, the data for this survey will be collected using web-based survey tool named Webropol. The survey will be shared through author's social media platforms which are LinkedIn, Facebook and Instagram.

5.4 Data analysis method

Common error in conducting quantitative study is that one does not think the data analysing phase until the information is gathered or the survey has been done. Usually, the results of a survey can be concluded in either of two ways: the researcher should learn the formula for each technique and apply the data to it or; use computer software to analyse the data. Analysing the data with a software is more beneficial as it imitates the modern way actual data analysis is done (Bryman, 2012). In order for the author to analyse the survey's results in numerical and sophisticated way, the author shall use IBM SPSS, which is one of the most used software for analysing quantitative data.

5.5 Scale of measurement

According to Hair et al. (2017) measurement scale "is a tool with a predetermined number of closed-ended responses that can be used to obtain an answer to a question" (Hair et al., 2017, p. 22). Measurement scales can be shared into four: nominal, ordinal, interval and ratio, which each represent a different level of measurement.

Nominal scale, also known as categorical scale, predicts numbers used to classify attributes such as people, professions or products. Nominal scales can include two or more Primer on Partial Least Squares categories. In case of multiple categories, each category has to be mutually exclusive and all other possible categories need to be included. A certain number can be pointed to identify each category. The numbers can be used to count the number of responses or percentage in all categories. Due to the nature of nominal scale, it is also the most restrictive scale.

Ordinal scale measures important information received by following the change of a value of a variable. When a variable is measured in ordinal scale, the increase or decrease in the variable's value predicts whether it is significant or not (Hair et al., 2017). The answers from this kind of survey can be placed on a continuum with the belief that some categories will exceed others. In ordinal scale, the difference between certain categories cannot be measured (Dawson, 2002).

Interval scale predicts accurate information about the rank order at which an observation is measured. The value used in an interval scale can be essentially any kind of

mathematical value such as mean or standard deviation. Exact comparison can be executed between these scales (Hair et al., 2017). Some examples of the use of an interval scale can be questions concerning age or household income (Dawson, 2002).

The scale that implements the most information is the ratio scale. For example, the value of 0 of a variable means that the variable is lacking from the observation. Ratio scale can be used when measuring i.e. length, volume or time (Hair et al., 2017).

The scale used to measure the items within the thesis survey is the Likert scale; which items are supported by the principles of an interval and nominal scale. The Likert scale was developed in 1932 by Rensis Likert in order to observe people's attitudes (Likert, 1932). Ever since the Likert scale was established, it has been widely used in numerous surveys, which measure the importance of different attitudes. The scale has been used for example in IS research (Bellman et al., 2004; Heirman et al., 2013; McGill, 2004; Ng et al., 2009).

5.6 Sample

The questionnaire link was shared with over 500 potential respondents on Facebook and over 100 contacts via Linkedin. Facebook and Linkedin connections were able to share the link forward to their networks (Snowballing technique). Sample size is recommended to be ten times the highest number of structural paths to a latent variable according to Hair et al. (2011). The sample size in this thesis is thus 110.

5.7 Questionnaire

The questionnaire consists of two parts. In the first part, demographic data such as age, gender, education, device use and general knowledge about data security will be asked. The second part of the questionnaire deals with questions related to each latent variable, which will predict their relations towards data security behaviour and perceived seriousness' relation to all the other variables. The seven-point Likert scale will be used in the questions.

6 DATA ANALYSIS

The purpose of this chapter is to perform analysis of the collected data. By utilizing two-step PLS-SEM analysis with a software called Smart-PLS, the author executes a model evaluation of the conceptual model presented in chapter 4.2. First, the outer model is observed by examining the outer loadings of the measurement items (i.e. measurement model). Thus, the reliability and validity of the key constructs can be ensured. Next, the inner model will be analysed so that the research hypotheses (i.e. conceptual model) can be tested.

6.1 Descriptive statistics

The responses were collected by using the Webropol platform. The responses were collected between 8th of April 2019 and 1st of May 2019. 133 respondents participated in the survey and all of them answered every question in the survey, so there were not uncompleted responses. Microsoft Excel was used to demonstrate the sample of the research. In the beginning of the survey, respondents were asked to state demographic information about themselves, which are illustrated below.

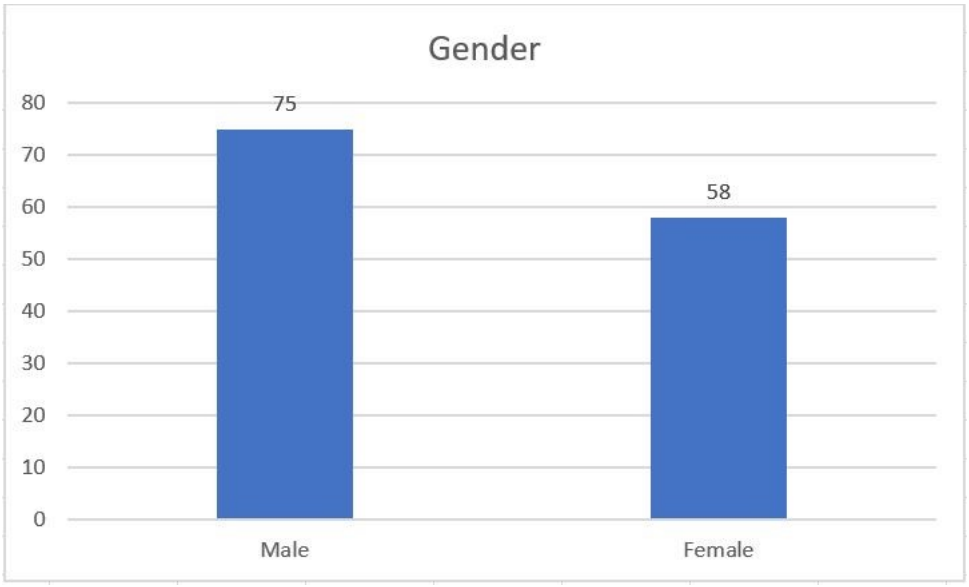

Figure 11. Gender of the respondents

Figure 11 shows that out of total respondents of 133, the number of male respondents was 75 and female respondents 58. The gender of respondents was distributed so that 56.39 percent were male, and 43.61 percent were female.

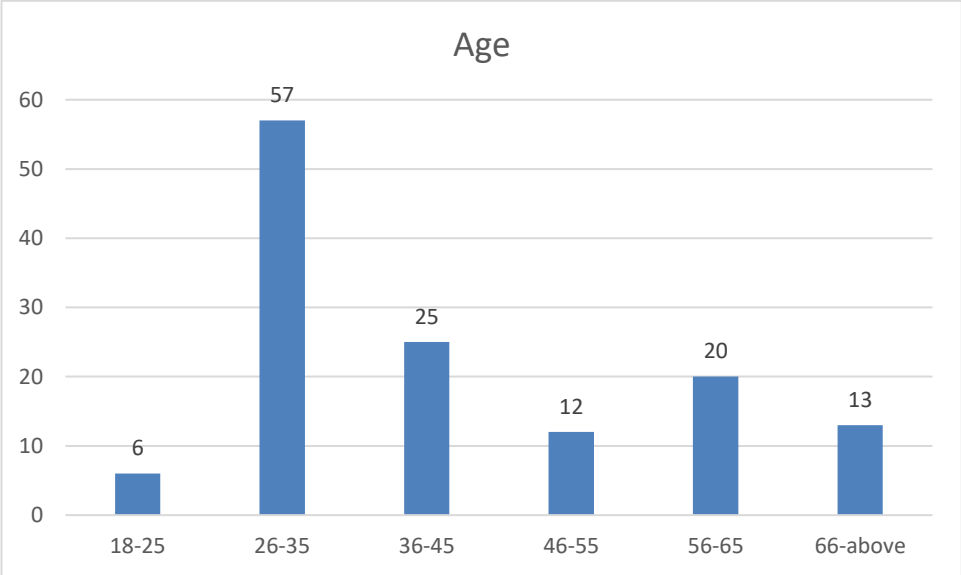

Figure 12. Age of the respondents

Figure 12 shows the age of respondents. Some variation can be detected: most of the respondents, 42.86 percent were between 26 to 35 years old. From the figure it can also be detected that 18.8 percent of the respondents were between 36 to 45 years old, 15.04 percent were between 56 to 65 years old, 9.77 percent were 66 years of age or older, 9.02 percent were between 46 to 55 years old and 4.51 percent were between 18 to 25 years old.

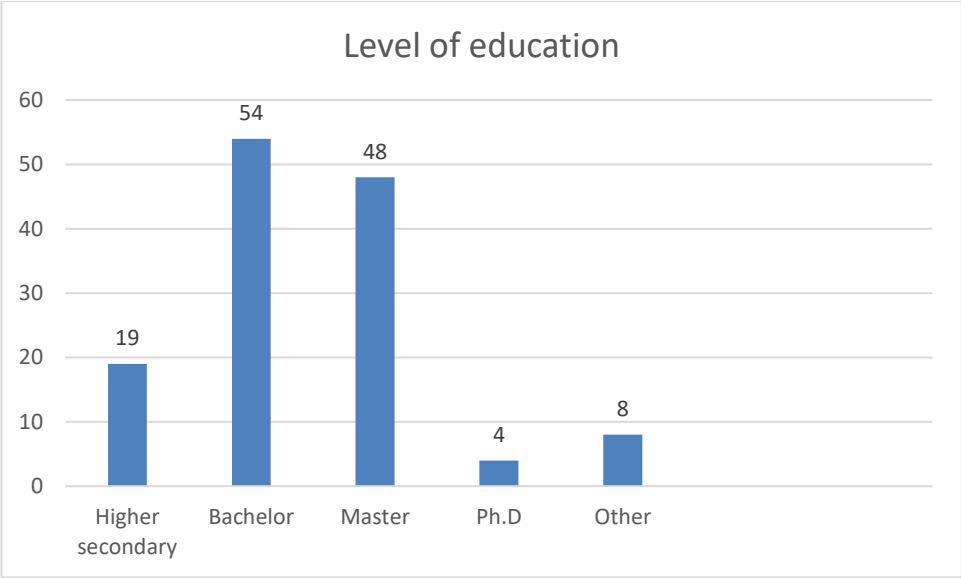

Figure 13. Level of education of the respondents

Figure 13 shows the level of education of the respondents. Most of the respondents, 40.60 percent obtained bachelor's degree. 36.09 percent of the respondents had master's degree, 14.29 percent had higher secondary certificate, other education 6.02 percent and 3.00 percent had Ph.D. degree.

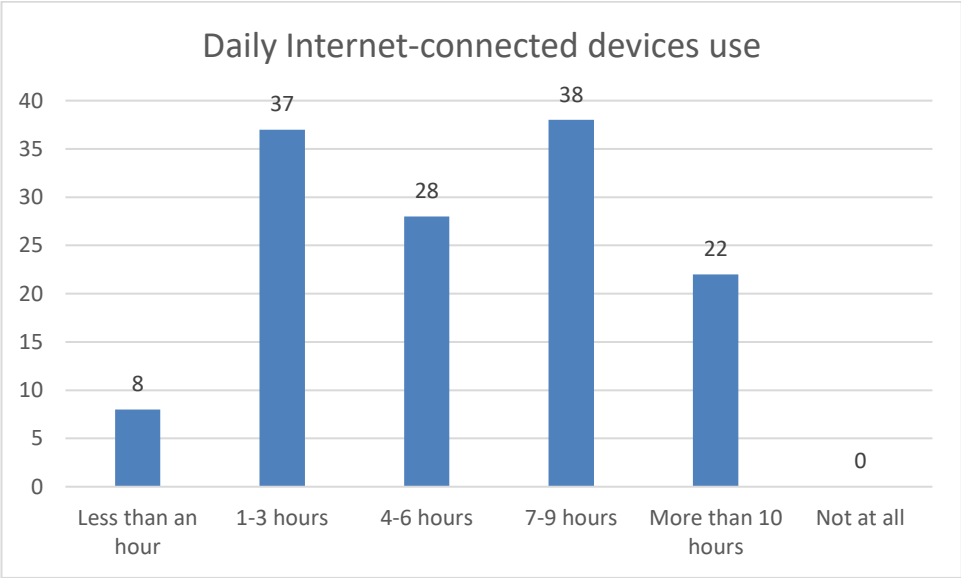

Figure 14. Daily Internet-connected devices use of the respondents

Figure 14 shows how many hours the respondents used internet-connected devices daily. Here it can be noticed that all respondents use internet-connected devices. 28.57 percent of respondents use internet-connected devices seven to nine hours a day, 27.82 percent of respondents one to three hours a day, 21.05 percent of respondents four to six hours a

day, 16.54 percent of respondents more than 10 hours a day and 6.02 percent of respondents less than an hour a day use the internet.

6.2 Outer model analysis

To start with the analysis, the reliability and validity of the outer model will be assessed. Foremost, the factor loadings of the model are calculated for each construct of the model to identify the validity and reliability of the outer model. According to Hair et al. (2014), the loading is recommended to be 0.70 or higher to point out that the key construct describes over 50 percent of the measurement item's variance. Composite reliability measure will be used to evaluate the internal consistency reliability. This thesis will be conducted so that the composite reliability should be 0.60 or higher as in exploratory research all values between 0.60 and 0.70 can also be accepted (Hair et al., 2011). Thus, the higher the values of composite reliability, the higher the internal consistency reliability will be. Following the outer model analysis, next the internal consistency of reliability will be assessed. In order to conduct this, Cronbach's alpha and composite reliability will be observed. Hair et al. (2014) suggest that composite reliability brings better results when assessing internal consistency reliability.

Next, the convergent validity of each construct will be assessed. Convergent validity is measured in order to see, to which extent the key construct coincides with its measurement items by measuring variance. This will be done by measuring the Average Variance Extract (AVE) for each construct. Hair et al. (2014) suggest that the value of AVE should exceed 0.50, as it indicates that over 50 percent of the variance is explained by the key constructs which the measurement items are related (Hair et al., 2014).

After the validity and reliability are stablished, the discriminant validity of the key constructs are assessed. The discriminant validity can be established by using Fornell Larcker criterion. The Fornell Larcker criterion shows, to what extent the constructs in the model correlate with each other by comparing the constructs' AVE value. Hair et al. (2014) suggest that constructs should not share variance in the model greater than the construct's own AVE value. Discriminant validity can also be observed by examining the cross loadings of the key measurement items in the model. In this case, the loadings of an indicator should exceed all its cross loadings (Hair et al., 2014).

6.2.1 Results analysis

According to Hair et al. (2011), measurement items which do not have acceptable characteristics while measuring reliability and validity should be removed from the conceptual model. While performing the analysis, measurement item SE1 showed low loading of 0.207 (<0.60), measurement item SE3 loading was 0.111 (<0.60), measurement item SE4 loading was 0.340 (<0.60), measurement item CA1 loading was 0.494 (<0.60), measurement item PB1 loading was 0.330(<0.60) and measurement item PB2 loading was 0.417(<0.60). As the AVE for CA was 0.508 (>0.50), the low loadings for items CA1, CA4 and CA5 were ignored. The AVE for Self-Efficacy was 0.561 (>0.50) improved after removing measurement items SE3 and SE4, as well as measurement items PB1 and PB2. After removing the items and running the analysis once more, all constructs' AVE values were higher than 0.50. Composite variability values for key constructs were significant as the lowest value was 0.817, which was higher than 0.70. Cronbach's alpha values for all constructs was above the recommended threshold, apart from Perceived Benefits (0.557) which is lower than 0.70. Nevertheless, Hair et al. (2014) state that Cronbach's alpha tends to undermine internal consistency reliability and thus the results rely more on composite reliability values.

Table 1 shows the results of the reliability and validity. Overruled items are such that were below the recommended threshold value and thus were removed from the analysis. Highlighted values are values that have been affected significantly by removing items which were not fitted with the recommended cut-off values. Values that are not highlighted stay significantly unaffected.

Table 1. Validity and reliability

Key Construct	Item	Outer loading	VIF (<5.0)	Cronbach's Alpha	CR	(AVE)
Cues to Action	CA1	0.494	1.046	0.752	0.834	0.508
	CA2	0.832	2.616			
	CA3	0.810	2.460			
	CA4	0.698	2.461			
	CA5	0.679	2.362			
Perceived Barriers	PBR1	0.870	2.143	0.820	0.868	0.625
	PBR2	0.869	1.857			
	PBR3	0.706	1.405			
	PBR4	0.701 0.704	1.923			
Perceived Benefits	~~PB1~~	0.330	1.783	0.557	0.817	0.691
	~~PB2~~	0.417	1.771			
	PB3	0.794 0.793	**(1.238)** 1.175			
	PB4	0.814	**(1.222)** 1.175			
Perceived Seriousness	PSE1	0.876	2.364	0.751	0.860	0.675
	PSE2	0.873 0.867	2.342			
	PSE3	0.694 0.708	1.193			
Perceived susceptibility	PSS1	0.838	2.118	0.843	0.895	0.680
	PSS2	0.827	2.025			
	PSS3	0.820	2.021			
	PSS4	0.814	2.067			
Data Privacy Behaviour	BEH1	0.967	4.328	0.934	0.968	0.938
	BEH2	0.970	4.328			
Self-efficacy	SE1	0.353	1.106	0.473	0.676	0.561
	SE2	0.999	**(1.121)** 1.106			
	~~SE3~~	0.111	1.011			
	~~SE4~~	-0.340	1.076			

AVE: Average Variance Extracted
CR: Composite Reliability

After establishing the validity and reliability of the outer model, discriminant validity needs to be observed according to the Fornell and Larcker criterion. As already stated, the observation can be conducted by comparing each key construct's AVE value with all the other key constructs in the model. A key construct should not share variance with other key constructs in the model that is greater than its own AVE value. Discriminant validity is observed so that it can be confirmed that constructs which are not supposed to

be related, truly remain that way. Table 2 shows that indeed, in this model the AVE values of the key constructs are diagonally higher on themselves than with other constructs.

Table 2. Discriminant Validity

	CA	PBR	PB	PSE	PSS	BEH	SE
CA	**0.713**						
PBR	0.023	**0.791**					
PB	0.146	0.230	**0.831**				
PSE	0.349	0.149	0.241	**0.821**			
PSS	0.180	0.193	0.254	0.530	**0.825**		
BEH	0.182	0.095	0.130	0.318	0.447	**0.969**	
SE	-0.110	-0.099	-0.035	-0.328	-0.316	-0.155	**0.749**

As mentioned previously, discriminant validity can also be observed by examining the cross loadings of the key measurement items in the model. Hair et al. (2011) suggest that each measurement item should have higher loading on its own key construct than on any other key construct. The cross-loadings table in Appendix 1 shows that this qualification is met in this study. It can be stated that the discriminant validity meets the requirements both with the Fornell and Larcker criterion as well as cross loading examination.

6.3 Inner model analysis

After the outer model analysis has been conducted, the next step in the PLS-SEM analysis is to assess the inner model. Urbach and Ahlemann (2010) suggest that significant level of path coefficient should be at least 0.5. Literature also supports the significant level of 0.10 (Wong, 2013) or even 0.05 (Lohmöller, 1989). By using the bootstrapping method provided by SmartPLS, the significance of path coefficients can be detected. Also, the research hypotheses can be tested systematically with bootstrapping. Bootstrapping is based on resampling subsamples of given data samples. Garson (2016) presents using minimum of 5000 subsamples to receive exact results. Thus, the sample used in this thesis is 5000. *t*-statistics and *p* values can be obtained while conducting bootstrapping, which are needed to test the significance of path coefficients. Hair et al. (2011) propose that the value of *t*-statistics should be 2.58% at 1% level of significance, 1.96 at 5% level of significance and 1.65 at 10% level of significance. To continue, Garson (2016) states that *p* value <0.05 is considered significant.

When detecting the variance in the model, R^2 value needs to be observed. R^2 value explains the key indicators and how they are connected to the dependent variable. R^2 value 0.20 or higher is seen high in consumer behaviour research. In IS research, R^2 value >0.49 is acceptable but at the same time, values 0.75, 0.50 and 0.23 can be stated to be not that significant in marketing research. (Hair et al., 2011) According to Urbach and Ahlemann (2010), the coefficient of determination is known as R^2 which is used to analyse the explained variance of a variable which is related to overall variance. In another word, R^2 is used to identify how solid the predicting powerful a latent construct based on its independent variable is. Chin (1998) described the significant levels of R^2 which are given below in Table 3. If R^2 is higher than 0.67 which means, the predicting power is substantial. If R^2 is higher than 0.33 which means, the predicting power is moderate. If R^2 is higher than 0.19, shows the value that the predicting power is weak.

Table 3. R^2 values

	R Square	R Square Adjusted
Cues to Action	0.121	0.115
Perceived Barriers	0.022	0.015
Perceived Benefits	0.058	0.051
Perceived susceptibility	0.281	0.275
Safe Data Privacy Behaviour	0.215	0.178
Self-efficacy	0.108	0.101

6.3.1 Results analysis and hypotheses testing

As mentioned in the earlier chapter, the significance of hypotheses can be detected by using inner model analysis. The following format explains the results: (Hx, β, *t, p*). in which Hx = tested hypotheses, β = path coefficient, *t* = t-statistics and *p* = p value. The hypotheses tested in this thesis are listed below:

H1. Perceived susceptibility to security incidents is positively related to data security behaviour.
H2. Perceived benefits of practicing data security measures are positively related to data security behaviour.
H3: Perceived barriers of practicing data security measures are negatively related to data security behaviour.
H4: Cues to action are positively related to data security behaviour.
H5: Self-efficacy is positively related to data security behaviour.

H6: Perceived seriousness of security incidents is positively related to data security behaviour.
H6a: Perceived seriousness of security incidents increases the positive effect of perceived susceptibility on data security behaviour.
H6b. Perceived seriousness of security incidents reduces the positive effect of perceived benefits on data security behaviour.
H6c. Perceived seriousness of security incidents reduces the negative effect of perceived barriers on data security behaviour.
H6d. Perceived seriousness of security incidents increases the positive effect of cues to action on data security behaviour.
H6e. Perceived seriousness of security incidents reduces the positive effect of self-efficacy on data security behaviour.

Table 4. Hypotheses testing

	β	***t*-Statistics**	***p***
Perceived susceptibility -> Data Privacy Behaviour	0.388	4.128	0.000***
Perceived Benefits -> Data Privacy Behaviour	-0.003	0.036	0.971
Perceived Barriers -> Data Privacy Behaviour	0.007	0.062	0.950
Cues to Action -> Data Privacy Behaviour	0.083	0.923	0.356
Self-efficacy -> Data Privacy Behaviour	0.006	0.068	0.946
Perceived Seriousness -> Data Privacy Behaviour	0.085	0.856	0.392
Perceived Seriousness -> Perceived susceptibility	0.530	8.520	0.000***
Perceived Seriousness -> Perceived Benefits	0.241	2.601	0.010**
Perceived Seriousness -> Perceived Barriers	0.149	1.047	0.296
Perceived Seriousness -> Cues to Action	0.349	4.208	0.000***
Perceived Seriousness -> Self-efficacy	-0.328	3.043	0.002**

Note: P<0.05= * P<0.01= ** P<0.001= ***

Based on the outcome of hypotheses testing (Table 4), it can be stated that H1 is supported by the model (H1, $\beta = 0.388$, $t = 4.128$, $p < 0.001$), which indicates that perceived susceptibility towards security incidents is positively related to data security behaviour and is statistically significant. The second hypothesis is not supported by the model (H2, $\beta = -0.003$, $t = 0.036$, $p = 0.971$) and the same goes with the third hypothesis (H3, $\beta = 0.007$, $t = 0.062$, $p = 0.950$). Also, hypotheses H4, H5 and H6 did not indicate any significance (H4, $\beta = 0.083$, $t = 0.923$, $p = 0.356$), (H5, $\beta = 0.006$, $t = 0.068$, $p = 0.946$), (H6, $\beta = 0.085$, $t = 0.856$, $p = 0.392$). Hypotheses H6a is supported and illustrates statistical significance (H6a, $\beta = 0.530$, $t = 8.520$, $p < 0.001$) as well as hypothesis H6b

(H6b, β = 0.241, *t* = 2.601, *p* = 0.010). Hypothesis H6c is rejected (H6c, β = 0.149, *t* = 1.047, *p* = 0.296). However, hypotheses H6d and H6e are supported by the model (H6d, β = 0.349, *t* = 4.208, *p* < 0.001) (H6e, β = -0.328, *t* = 3.043, *p* < 0.01), where H6d is statistically very significant and H6e shows weak significance. In total, out of 11 hypotheses, 6 were rejected. Thus, the remaining 5 hypotheses show statistical significance and supported by the model.

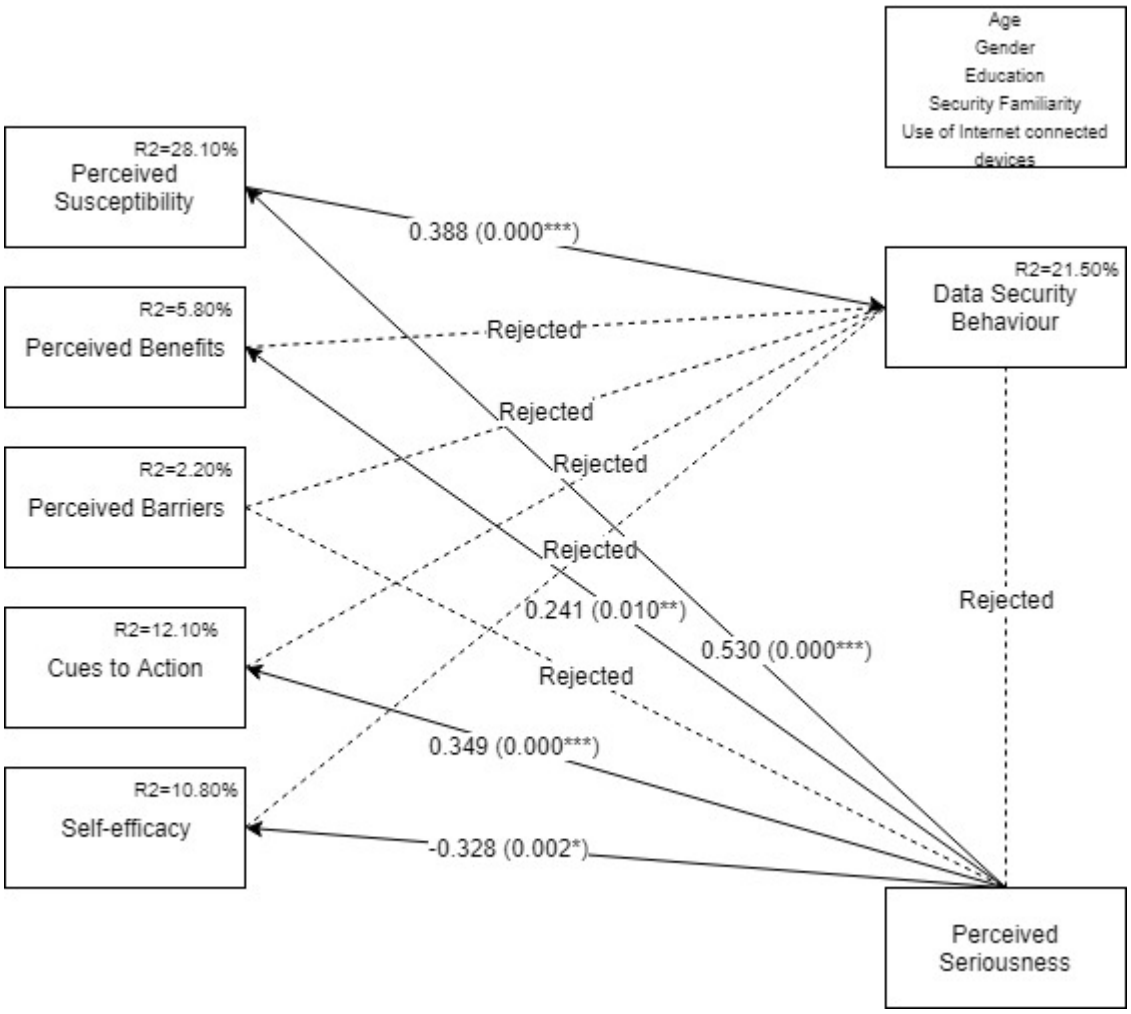

Figure 10. Conceptual model results
Note: P<0.05= * P<0.01= ** P<0.001= ***

6.4 Multi-Group Analysis

Usually PLS path modelling is placed on the expectation that the analysed data is based on a single population. Yet, this kind of assumption of homogeneity is impractical, when thinking of the notions the real-world sets, as individuals tend to be heterogeneous in their perceptions and evaluations. (Sarstedt et al., 2011) According to Sarstedt et al. (2011),

multi-group analysis (MGA) can be used to detect the probability of population differing from sub-populations.

In the questionnaire, respondents were asked to provide information concerning their demographic data (age, gender, education, device use and general knowledge about data security). Based on the received responses, respondents could be divided into different groups. Firstly, the author will assess the difference between genders in different paths (see Table 5).

The path between Perceived Susceptibility and Data Security Behaviour can be deemed as significant for both female (β = 0.564; t = 2.931; p = 0.004) and male (β = 0.309; t = 0.102; p = 0.006). Then again, the path relation between Perceived Benefits and Data Security Behaviour did not indicate significance on either gender. For female the values showed (β = -0.048; t = 0.386; p = 0.700) and for male (β = 0.016; t = 0.107; p = 0.915). The path relation between Perceived Barriers and Data Security Behaviour indicated that neither female (β = 0.095; t = 0.591; p = 0.554) nor male (β = -0.004; t = 0.026; p = 0.979) showed statistical significance. The path relation between Cues to Action and Data Security Behaviour illustrated that for female (β = -0.036; t = 0.304; p = 0.761) or male (β = 0.152; t = 0.917; p = 0.360) there is not any statistical significance. The path relation between Self-efficacy and Data Security Behaviour showed that for female (β = 0.121; t = 0.827; p = 0.409) or male (β = -0.078; t = 0.519; p = 0.604) there is no significance. Again, in the path relation between Perceived Seriousness and Data Security Behaviour, the female (β = 0.079; t = 0.489; p = 0.625) and male (β = 0.077; t = 0.550; p = 0.582) did not indicate any statistical significance. The path relation between Perceived Seriousness and Perceived Susceptibility showed that for female there can be found statistical significance (β = 0.686; t = 11.096; p = 0.000), whereas for male (β = 0.417; t = 4.086; p = 0.000), significance can also be detected. The path relation between Perceived Seriousness and Perceived Benefits shows significance for female (β = 0.386; t = 3.315; p = 0.001) but not for male (β = 0.127; t = 0.859; p = 0.391). The path relation between Perceived Seriousness and Perceived Barriers shows no significance for female (β = 0.036; t = 0.172; p = 0.864) or for male (β = 0.219; t = 1.243; p = 0.214). The path relation between Perceived Seriousness and Cues to Action illustrated significance for the female (β = 0.469; t = 4.998; p = 0.000) but not for male (β = 0.228; t = 1.132; p = 0.258). The path relation between Perceived Seriousness and Self-efficacy was significant for female (β = -0.455; t = 3.608;

p = 0.000) but not for male (β = -0.274; t = 0.958; p = 0.339). The most significant difference between female and male was shown to be between Perceived Seriousness and Perceived Susceptibility, where the p value was 0.013.

Table 5. Multi-group Analysis for Gender

	β (Female)	β (Male)	*t*-statistics (Female)	*t*-statistics (Male)	*p*-(Female)	*p*-(Male)
Cues to Action -> Safe Data Privacy Behaviour	-0.036	0.152	0.304	0.917	0.761	0.360
Perceived Barriers -> Safe Data Privacy Behaviour	0.095	-0.004	0.591	0.026	0.554	0.979
Perceived Benefits -> Safe Data Privacy Behaviour	-0.048	0.016	0.386	0.107	0.700	0.915
Perceived Seriousness -> Cues to Action	0.469	0.228	4.998	1.132	0.000	0.258
Perceived Seriousness -> Perceived Barriers	0.036	0.219	0.172	1.243	0.864	0.214
Perceived Seriousness -> Perceived Benefits	0.386	0.127	3.315	0.859	0.001	0.391
Perceived Seriousness -> Perceived susceptibility	0.686	0.417	11.096	4.086	0.000	0.000
Perceived Seriousness -> Safe Data Privacy Behaviour	0.079	0.077	0.489	0.550	0.625	0.582
Perceived Seriousness -> Self-efficacy	-0.455	-0.274	3.608	0.958	0.000	0.339
Perceived susceptibility -> Safe Data Privacy Behaviour	0.564	0.309	2.931	2.759	0.004	0.006
Self-efficacy -> Safe Data Privacy Behaviour	0.121	-0.078	0.827	0.519	0.409	0.604

The next MGA was conducted based on age (Table 6). The age of the respondents was divided into six groups (aged between 18-25, 26-35, 36-45, 46-55, 56-65, 66-above). The groups have been divided into two, so that in the first group will be those respondents who were between 18-45 years old and in the second group those who were from 46 to above 66 years old. The Table 6 shows differences in the age groups in different paths.

Table 6. Multi-Group Analysis for Age

	β (Age_1.0)	β (Age_ 2.0)	*t*- (Age_ 1.0)	*t*- (Age_ 2.0)	*p*- (Age_1.0)	*p*- (Age_ 2.0)
Cues to Action -> Safe Data Privacy Behaviour	-0.021	0.118	0.165	0.657	0.869	0.511
Perceived Barriers -> Safe Data Privacy Behaviour	0.070	0.056	0.616	0.319	0.538	0.750
Perceived Benefits -> Safe Data Privacy Behaviour	-0.003	-0.028	0.032	0.131	0.975	0.895
Perceived Seriousness -> Cues to Action	0.358	0.528	2.669	3.903	0.008	0.000
Perceived Seriousness -> Perceived Barriers	0.279	-0.129	2.107	0.579	0.036	0.563
Perceived Seriousness -> Perceived Benefits	0.334	0.210	2.784	1.005	0.006	0.315
Perceived Seriousness -> Perceived susceptibility	0.552	0.487	6.419	4.326	0.000	0.000
Perceived Seriousness -> Safe Data Privacy Behaviour	0.120	0.135	0.903	0.683	0.367	0.495
Perceived Seriousness -> Self-efficacy	-0.401	-0.292	2.894	1.821	0.004	0.069
Perceived susceptibility -> Safe Data Privacy Behaviour	0.315	0.441	2.399	3.129	0.017	0.002
Self-efficacy -> Safe Data Privacy Behaviour	-0.142	0.089	1.062	0.483	0.289	0.629

The path between Perceived Susceptibility and Data Security Behaviour can be deemed as significant for both age groups 1 (β = 0.315; *t* = 2.399; *p* = 0.017) and 2 (β = 0.441; *t* = 3.129; *p* = 0.002). The path relation between Perceived Benefits and Data Security Behaviour did not indicate significance on either age group. For group 1 the values showed (β = -0.003; *t* = 0.032; *p* = 0.975) and for group 2 (β = -0.028; *t* = 0.131; *p* = 0.895). The path relation between Perceived Barriers and Data Security Behaviour indicated that neither group 1 (β = 0.070; *t* = 0.616; *p* = 0.538) nor group 2 (β = 0.056; *t* = 0.319; *p* = 0.750) showed statistical significance. The path relation between Cues to Action and Data Security Behaviour illustrated that for group 1 (β = -0.021; *t* = 0.165; *p* = 0.869) or group 2 (β = 0.118; *t* = 0.657; *p* = 0.511) there is not any statistical significance. The path relation between Self-efficacy and Data Security Behaviour showed that for group 1 (β = -0.142; *t* = 1.062; *p* = 0.289) or group 2 (β = 0.089; *t* = 0.483; *p* = 0.629) there is no significance. In the path relation between Perceived Seriousness and Data Security Behaviour, the group 1 (β = 0.120; *t* = 0.903; *p* = 0.367) and group 2 (β = 0.077; *t* = 0.550; *p* = 0.582) did not indicate any statistical significance. The path

relation between Perceived Seriousness and Perceived Susceptibility showed that for group 1 there can be found statistical significance (β = 0.552; t = 6.419; p = 0.000), as well as for group 2 (β = 0.487; t = 4.326; p = 0.000). The path relation between Perceived Seriousness and Perceived Benefits shows significance for group 1 (β = 0.334; t = 3.315; p = 0.001) but not for group 2 (β = 0.210; t = 1.005; p = 0.315). The path relation between Perceived Seriousness and Perceived Barriers showed significance for group 1 (β = 0.279; t = 2.107; p = 0.036) but not for group 2 (β =-0.129; t = 0.579; p = 0.563). The path relation between Perceived Seriousness and Cues to Action illustrated significance for group 1 (β = 0.358; t = 2.669; p = 0.008) and for group 2 (β = 0.528; t = 3.903; p = 0.000). The path relation between Perceived Seriousness and Self-efficacy was significant for group 1 (β = -0.401; t = 2.894; p = 0.004) but not for group 2 (β = -0.292; t = 1.821; p = 0.069). The most significant difference between group 1 and group 2 was shown to be between Perceived Seriousness and Perceived Barriers, where the p value was 0.063. Although the $p > 0.50$, it still brings some insight considering the difference between the age groups.

Thirdly, the author shall conduct MGA based on the respondents use of Internet-connected devices. The responses were divided into two groups based on the time they used Internet-connected devices per day. The groups are light users, who use devices between less than an hour up to six hours a day and heavy users, who use devices between 7 hours or to more (Table 7).

Table 7. Multi-Group Analysis for Daily Internet-connected Device Use

	β (Light users)	β (Heavy users)	*t*- (Light users)	*t*- (Heavy users)	*p*- (Light users)	*p*- (Heavy users)
Cues to Action -> Safe Data Privacy Behaviour	0.162	-0.008	1.426	0.044	0.155	0.965
Perceived Barriers -> Safe Data Privacy Behaviour	0.067	-0.067	0.502	0.374	0.616	0.708
Perceived Benefits -> Safe Data Privacy Behaviour	0.096	-0.207	0.886	1.222	0.376	0.222
Perceived Seriousness -> Cues to Action	0.442	0.350	3.696	1.672	0.000	0.095
Perceived Seriousness -> Perceived Barriers	0.015	0.259	0.077	0.964	0.938	0.336
Perceived Seriousness -> Perceived Benefits	0.248	0.136	2.036	0.609	0.042	0.543
Perceived Seriousness -> Perceived susceptibility	0.579	0.358	6.415	3.324	0.000	0.001
Perceived Seriousness -> Safe Data Privacy Behaviour	0.089	-0.068	0.689	0.379	0.491	0.705
Perceived Seriousness -> Self-efficacy	-0.386	0.395	3.963	0.968	0.000	0.333
Perceived susceptibility -> Safe Data Privacy Behaviour	0.445	0.311	4.444	1.743	0.000	0.082
Self-efficacy -> Safe Data Privacy Behaviour	0.002	0.255	0.017	1.106	0.986	0.269

The path between Perceived Susceptibility and Data Security Behaviour can be stated as significant for light users (β = 0.445; t = 4.444; *p* = 0.000) but not for heavy users (β = 0.311; *t* = 1.743; *p* = 0.082). The path relation between Perceived Benefits and Data Security Behaviour did not indicate significance on either group. For light users the values showed (β = 0.096; *t* = 0.886; *p* = 0.376) and for heavy users (β = -0.207; *t* = 1.222; *p*=0.222). The path relation between Perceived Barriers and Data Security Behaviour indicated that neither light users (β = 0.067; *t* = 0.502; *p* = 0.616) nor heavy users (β = -0.067; *t* = 0.374; *p* = 0.708) showed statistical significance. The path relation between Cues to Action and Data Security Behaviour illustrated that for light users (β = 0.162; *t* = 1.426; *p* = 0.155) or heavy users (β = -0.008; *t* = 0.044; *p* = 0.965) there is not any statistical significance. The path relation between Self-efficacy and Data Security Behaviour showed that for light users (β = 0.002; *t* = 0.017; *p* = 0.986) or heavy users (β = 0.255; *t* = 1.106; *p* = 0.269) there is no significance. In the path relation between Perceived Seriousness and Data Security Behaviour, the light users (β = 0.089; *t* = 0.689; *p* = 0.491) and heavy users (β = -0.068; *t* = 0.379; *p* = 0.705) did not indicate any statistical significance. The path relation between Perceived Seriousness and Perceived Susceptibility showed that for light users there can be found statistical significance (β =

0.579; $t = 6.415$; $p = 0.000$), as well as for heavy users (β = 0.358; $t = 3.324$; $p = 0.001$). The path relation between Perceived Seriousness and Perceived Benefits shows significance for light users (β = 0.248; $t = 0.077$; $p = 0.042$) but not for heavy users (β = 0.136; $t = 0.609$; $p = 0.543$). The path relation between Perceived Seriousness and Perceived Barriers shows no significance for light users (β = 0.015; $t = 0.172$; $p = 0.938$) or for heavy users (β = 0.259; $t = 0.964$; $p = 0.336$). The path relation between Perceived Seriousness and Cues to Action illustrated significance for the light users (β = 0.442; $t = 3.696$; $p = 0.000$) but not for heavy users (β = 0.350; $t = 1.672$; $p = 0.095$). The path relation between Perceived Seriousness and Self-efficacy was significant for light users (β = -0.386; $t = 3.963$; $p = 0.000$) but not for heavy users (β = 0.395; $t = 0.968$; $p = 0.333$). The most significant difference between light users and heavy users can be said to be in the relation between Perceived Seriousness and Self-efficacy, where the p value was 0.984.

7 DISCUSSION AND RESULTS

This chapter discusses the findings of the thesis. The findings are presented with concluding remarks. Key findings are presented based on the analyses conducted in Chapter 6. After this, the author observes how well the research questions were answered in this thesis. The author shall also discuss the theoretical contributions and address practical implications.

7.1 Key findings

As discussed in Chapter 4, this thesis applies the Health Belief Model (HBM) as the theoretical basis for the research. Various studies have proven HBM to be applicable when studying behaviour related to computer security (e.g. Aytes et al., 2003; Ng et al., 2009). Per earlier introductions, the data security behaviour has not been studied as thoroughly. Traditionally, researches conducting HBM have proven Perceived Susceptibility, Perceived Benefits and Self-efficacy to be of importance in computer security context. When observing the conceptual model, the concepts illustrate weak predicting power for the constructs. However weak, according to the study Perceived Seriousness influences almost all the constructs. Perceived Seriousness of being afflicted to a security incident strengthens user's Perceived Susceptibility, in which the user can be said to be more aware of threats coming his way. Perceived Seriousness also affects to Perceived Benefits so that the user is willing to put possible additionally gained benefits aside in order to conduct countermeasures for his safety; security measures are followed even though the user might not fully see these measures as effective. When it comes to the Cues to Action, Perceived Seriousness has such an affect that the person will start to act according to safety measures as soon as a threat can be detected. Also, Perceived Seriousness and Self-Efficacy are in relation so that the user is likely to attempt practicing countermeasures, even though he might not be confident in his own skills but still seeks to try preventive measures in the fear of losing his information. In addition to the effects of Perceived Seriousness, also Perceived Susceptibility was found to have a relation with Data Security Behaviour. As stated earlier, people act differently and according to different behaviours. Even though same information is presented to different users, they can interpret them differently and make their decisions based on their perceptions. Yet, when greater susceptibility towards a security incident is detected, the user will likely act

according to greater level of data security behaviour. Surprisingly, the results of this thesis did not indicate statistical significance in the relations of Self-Efficacy and Data Security Behaviour and Perceived Benefits and Data Security Behaviour. According to the proposed model, they have significance when observed from the Perceived Seriousness point of view and when observed with the help of MGA.

When interpreting the results of MGA, it can be noted that both male and female feel that the possibility of having their personal information threatened effects on their behaviours. Female respondents show more willingness to use countermeasures, although they might not be sure of the effectiveness of them. Female respondents also seem to act according to safe behaviour as soon as they detect a threat and despite feeling to be fully capable doing it, whereas male do not.

MGA also brought up differences in different age groups. All age groups showed that when they suspect that there is a possibility for a security incidence, more countermeasures are performed. This can be observed in responses from the respondents aged between 18-45 years who illustrates that they will conduct countermeasures despite not being entirely sure of their effectiveness but just to have some kind of protection. This age group was also more willing to overcome their barriers in practicing safety measures. Additionally, age group 18-45 years old were willing to conduct safety measures whether they trusted their capabilities or not.

When comparing the light users and heavy users of the Internet, the light users showed more statistical significance in their analysis results. The analysis for light users showed that when they believe their safety is threatened, they are more likely to take action with countermeasures than the heavy users. Light users are also more likely to act according to safety measures as soon as they feel their security is afflicted. Also, light users showed more willingness towards conducting safety measures, even if they did not feel capable enough.

The proposed conceptual model proved to be impractical, as the constructs in the model did not show strong direct relations towards Data Security Behaviour. Thus, after testing the conceptual model, the model was revised. When altering the proposed model and observing it through indirect effects, more assuring values could be detected (Figure 15).

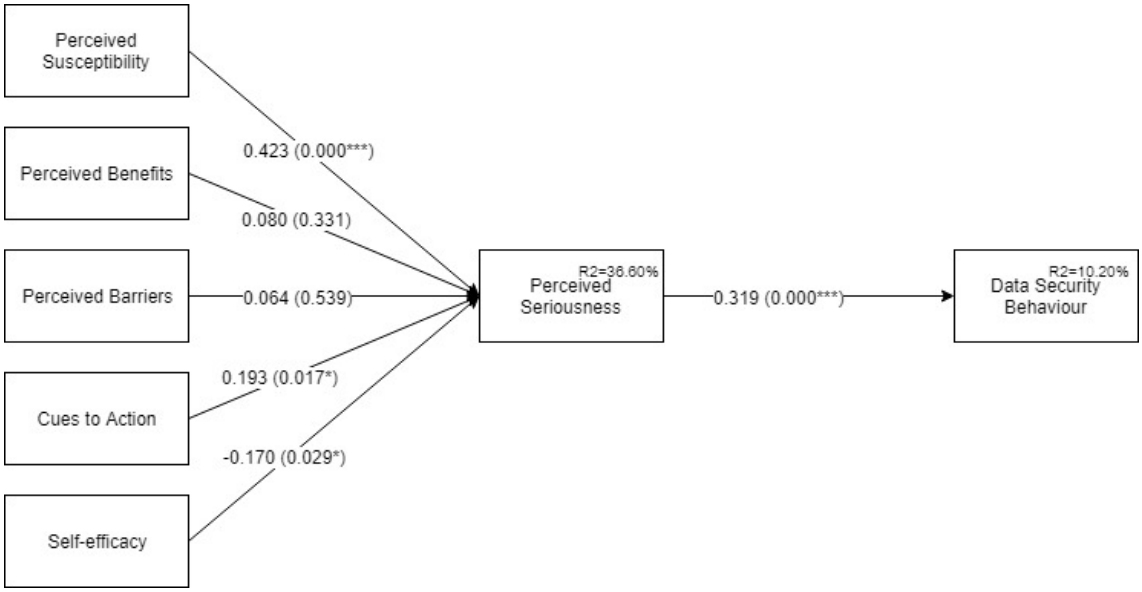

Figure 15. Revised conceptual model
Note: P<0.05= * P<0.01= ** P<0.001= ***

If the model would be altered so that Perceived Susceptibility, Perceived Benefits, Perceived Barriers, Cues to Action and Self-Efficacy would be argued to have an impact on Perceived Seriousness, which then again would have relation to Data Security Behaviour mediated by the Perceived Seriousness, the model shows more reliable results. So, the role of Perceived Seriousness would be important as a mediating variable towards Data Security Behaviour and the hypotheses could be postulated through the other constructs' relation towards Perceived Seriousness and thus its effect on Data Security Behaviour.

7.2 Research questions

The research questions are once more visited at this point of the discussion. The findings of this thesis and literature are also integrated to interpret the validity of the research questions of the thesis.

Research question 1: How does Perceived Seriousness of data privacy risk affect Data Privacy Behaviour?

The findings revealed, that even though Perceived Seriousness does not have direct linkage to Data Security Behaviour, it has an effect on majority of the constructs in the conceptual

model. According to the responses, Perceived Seriousness has an indirect relation to Data Privacy Behaviour via Perceived Susceptibility.

Research question 2: How do personal demographics affect perceptions concerning data behaviour?

The users aged 18-45 are more willing to overcome their barriers and take countermeasures in use, despite not fully trusting the effectiveness if the measures, or their own capabilities. All age groups are willing to improve their security behaviour, if they feel that their privacy is threatened. Female users are more likely to take measures in action and overcome their incapability. Then again, all users despite their gender are willing to take more countermeasures in action if their security is threatened. Also, users who use Internet-connected devices up to six hours a day show more willingness for using countermeasures if their security is afflicted.

Research question 3: Which constructs have a significant effect on data privacy behaviour?

According to the results, only one direct relation could be detected. Perceived Susceptibility has a direct linkage to Data Security Behaviour and Perceived Seriousness has an indirect relation when observed via Perceived Susceptibility. According to the R^2 values, the relations are not statistically significant but can yet be detected with the conceptual model.

7.3 Theoretical contributions

This thesis contributes to the Data Security literature by showing the findings on user's perception towards their own privacy and behaviour. Prior studies concerning computer security can be found from the literature but as data security is relatively new concern, mere researches concerning the subject can be found. Although the conceptual model proved to be impractical for the study's purpose, more insight towards user's perception was obtained. Thus, it can be said that users' perceptions vary depending on their background and perceived threat to their security. The findings of this thesis enrich the current Data Security literature so that in the future more studies from the field can be conducted.

7.4 Practical implications

In addition to the theoretical implications, this thesis also provides some practical implications considering user's behaviour regarding their data privacy. It is evident, based on the analysis that Internet-connected devices are used many hours a day and thus the risk of having your personal information being afflicted increases. A good number of respondents stated that they do not read thoroughly the terms and conditions of a service they are using. Considering renewed regulations concerning the user's rights, users should acknowledge the way their data is used and their rights over the data. Confidence towards using social media was also a factor, that rose in the answers; respondents feel that social media platforms are not trustworthy and simultaneously feel that their safety might be at risk. Thus, the service providers need to be as transparent as possible when it comes to asking for consent and ensuring proper security measures for their users. Additionally, many of the respondents answered that they do not always check the address or subject of an email they receive. Users need to pay more attention, which applications and websites they provide access and their personal information to.

8 CONCLUSION

The main objective of this thesis was to observe factors that have an influence on user's perception about their data security and which attributes effect on the behaviour, based on which they act accordingly. The Health Belief Model was used as a basis for the study, as well as previously conducted study based on computer security (Ng et al., 2009). The variables within the HBM including Perceived Susceptibility, Perceived Benefits, Perceived Barriers, Cues to Action, Self-Efficacy and Perceived Seriousness were used in order to find out the underlying beliefs towards one's data security behaviour.

This thesis has also some limitations. The results could have varied if the sample and the number of responses have been greater. When analysing the responses, the model showed weak predicting power. Once the model was altered, it brought more significance to direct linkages towards Data Security Behaviour. More indirect relations towards Data Security Behaviour also brought up more values, that could have been interpreted as significant outcome for the thesis. One reason for the weak predicting power can be found from the question setting. The questions did not support each other adequately and thus could have had an influence on the end result. Yet, the proposed questions were highly related to the modern era and everyday data security.

Securing one's information is a growing concern. The research in this thesis was conducted considering the users' daily activities and background, as well as the time spent with Internet-connected devices. In the field of IS, more study concerning user's behaviour towards their privacy in the modern era could be executed. Future studies may be conducted on the user's attitude and behaviour concerning threats rising from sharing one's information with service providers. As this study was conducted by using quantitative method, further study could follow qualitative method by collecting information from the respondents. Service provider's perspective is also a research objective that could be studied more thoroughly in the future.

REFERENCES

Afzal,W. & Thompson, K.M. (2011). Contributions of cognitive science to information science: An analytical synopsis. *Emporia State Research Studies,* **47**(1), pp. 18-23.

Ajzen, I. (1991) The Theory of Planned Behavior. *Organizational Behavior and Human Decision Process*, **50**, pp. 179-211.

Aytes, K., Connolly, T. (2003). A Research Model for Investigating Human Behavior Related to Computer Security. In *AMCIS 2003 Proceedings*.

Aytes, K., Connolly, T. (2004). Computer Security and Risky Computing Practices: A Rational Choice Perspective. *Advanced Topics in End User Computing***, 4**, pp. 257-279

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, 84(2), pp. 191-215

Bellman, S., Johnson, E.J., Kobrin, S.J., Lauder, J.H., Lohse G.L. (2004). International Differences in Information Privacy Concerns: A Global Survey of Consumers. *The Information Society*, **20**, pp. 313-324.

Benaroch, M., Chernobai, A. (2017). Operational IT failures, IT value destruction, and board-level IT governance changes. *MIS Quarterly*, **41**(3), pp. 729-762.

Bryman, A. (2012). *Social Research Methods*. Oxford University Press:US

Bulgurcu, B., Cavusoglu, H., Benbasat I. (2010). Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness. *MIS Quarterly*, **34**(3), pp. 523-548.

Chaffey, D. (2009). *E-Business and E-Commerce Management*. England: Pearson Education Limited.

Chen, D., Zhao, H. (2012). Data security and privacy protection issues in cloud computing. *2012 International Conference on Computer Science and Electronics Engineering*, **1**, pp. 647–651.

Choi, S.S., Choi, M. (2007). Consumer's Privacy Concerns and Willingness to Provide Personal Information in Location-Based Services. *The 9th International Conference on Advanced Communication Technology.*

Creswell, J. (2014). *Research Design – Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications Inc:US

Crossler, E., Johnston, A.C., Lowry, P.B., Hu, Q., Warkentin, M., Baskerville, R. (2002). Future directions for behavioral information security research. *Computers & Security*, **32** (2013), pp. 90-101.

Davis, F.D. Jr. (1985) *A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results*. Massachusetts: Massachusetts Institute of Technology.

Davis F.D, Bagozzi, R.P., & Warshaw, P.R. (1989). User Acceptance of Computer Technology: a Comparison of Two Theoretical Models. *Management Science*, **35** (8), pp. 982-1003.

Dawson, C. (2002). *Practical Research Methods.* UK: How To Books Ltd.

Dinev, T., Hu, Q. (2007). The Centrality of Awareness in the Formation of User Behavioral Intention toward Protective Information Technologies. *Journal of the Association for Information Systems*, **8**(7), pp. 386-408.

DLA Piper (2019). *GDPR Data Breach Survey*. DLA Piper.

Dodel, M., Mesch, G. (2017). Cyber-victimization preventive behavior: A health belief model approach. *Computers in Human Behavior***, 68**, pp. 359-367.

Dwivedi, H., Clark, C., Thiel, D. (2010). *Mobile Application Security*. The McGraw-Hill Companies.

European Union. (2018). *The European Union – What it is and what it does*. Luxembourg: Publications Office of the European Union.

European Union. (2018). *The EU in brief: From economic to political union*. Website of the European Union. Referred to January 13th, 2019. Available at: https://europa.eu/european-union/about-eu/eu-in-brief_en.

Fette, I., Sadeh, N., Tomasic, A. (2007). Learning to Detect Phishing Emails. In *WWW '07 Proceedings of the 16th international conference on World Wide Web*, pp. 649-656.

Fishbein, M. & Ajzen, I. (2010). *Predicting and Changing Behavior - The Reasoned Action Approach*. New York: Psychology Press.

Gao, X., Yang, Y., Fu, H., Lindqvist, J., Wang, Y. (2014). Private Browsing: an Inquiry on Usability. In *WPES '14 Proceedings of the 13th Workshop on Privacy in the Electronic Society*, pp. 97-106.

Garson, D. (2016). *Partial Least Squares: Regression & Structural Equation Models*. Statistical Associates Publishing.

Glanz, K., Rimer, B.K., Viswanath, K. (2010). *Health Education - Theory, Research, and Practice*. Jossey-Bass: San Francisco.

Gross, R., Acquisti, A. (2005). Information Revelation and Privacy in Online Social Networks. In *WPES '05 Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, pp. 71-80.

Hadlington, L. (2017). Human factors in cybersecurity; examining the link between Internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviours. *Heliyon*, 3 (2017).

Hair, J., F., Ringle, C., M. and Sarstedt, M. (2011) PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, **19**(2), pp. 139-152.

Hair, J. F., Ringle, C.M., Sarstedt, M., Smith, D., Reams, R. (2014). Partial least squares structural equation modelling (PLS-SEM): A useful tool for family business researchers. *Journal of Family Business Strategy*, **5**(1), pp. 105-115.

Hair, J.F. Jr., Hult, G.T.M., Ringle, C.M., Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM).* US: SAGE Publications, Inc

Hassel, L. & Wiedenbeck, S. (2004). Human Factors and Information Security. In *DIMACS Workshop on Usable Privacy and Security Software*, 7–8 July 2004, Piscataway, NJ.

Heinström, J. (2006). Psychological factors behind incidental information acquisition. *Library & Information Science Research*, **28**, pp. 579–594.

Heirman, W., Walrave, M., Ponnet, K. (2013). Predicting Adolescents' Disclosure of Personal Information in Exchange for Commercial Incentives: An Application of an Extended Theory of Planned Behavior. *Cyberpsychology, Behavior, and Social Networking*, **16** (2), pp. 81-87.

Isaak, J. & Hanna, M.J. (2018). User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer*, **51**(8), pp. 56-59.

Jayanti, R., K., Burns, A.C. (1998). The Antecedents of Preventive Health Care Behavior: An Empirical Study. *Journal of the Academy of Marketing Science*, **26**(1), pp. 6-15.

Johnson, J.D. & Case, D.O. (2012). Socio-Psychological Factors in Health; Models of Information Seeking. *Health Information Seeking*, 39-61, 96-122.

Johnson, J.D., Donahue, W.A., Atkin, C. K., Johnson, S. (1995). A Comprehensive Model of Information Seeking – Test Focusing on a Technical Organization. *Science Communication*, **16**(3), pp. 274-303.

Komatsu, A., Takagi, D., Takemura, T. (2013). Human aspects of information security - An empirical study of intentional versus actual behavior. *Information Management & Computer Security*, **21**(1), pp. 5-15.

Kothari, C.R. (2004). *Research Methodology*. New Delhi: New Age International Ltd.

LaRose, R., Rifon, N.J., Enbody, R. (2008). Promoting Personal Responsibility for Internet Safety. *Communications of the ACM*, **51**(3), pp. 71-76.

Lee, D., Larose, R., Rifon, N. (2008). Keeping our network safe: a model of online protection behaviour. *Behaviour & Information Technology*, **27**(5), pp. 445-454.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, **22**(140).

Lohmöller, J. (1989). *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg: Physica-Verlag.

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

McGill, T. (2004). The Effect of End User Development on End User Success. *Advanced Topics in End User Computing*, **4**, pp. 21-41.

National Institute of Standards and Technology. (2013). *Glossary of Key Information Security Terms.* National Institute of Standards and Technology.

Ng, B., Kankanhalli, A., Xu, Y. (2009). Studying users' computer security behavior: A health belief perspective. *Decision Support Systems,* **46**, pp. 815–825.

Organisation for Economic Co-operation and Development. (2007). *Glossary of Statistical Terms*. Organisation for Economic Co-operation and Development.

Prochaska, J., DiClemente, C. (1992). *The Transtheoretical Approach. Handbook of Psychotherapy Integration*. New York: Basic Books.

Rao, R. V., Selvamani, K. (2015). Data Security Challenges and Its Solutions in Cloud Computing. *Procedia Computer Science***, 48**, pp. 204-209.

Rogers, L. (2006). *Home Computer Security*. CERT Coordination Centre.

Rosenstock, I.M. (1974). Historical Origins of the Health Belief Model. *Health Education Monographs,* **2**(4), pp. 328-335.

Sarstedt, M., Henseler, J., Ringle, C. M. (2011). Multigroup Analysis in Partial Least Squares (Pls) Path Modeling: Alternative Methods and Empirical Results. *Measurement and Research Methods in International Marketing (Advances in International Marketing),* **22**, pp. 195–218.

Savolainen, R. (2006). Spatial factors as contextual qualifiers of information seeking. *Information Research*, **11**(4).

Steinfeld, N. (2015). "I agree to the terms and conditions": (How) do users read privacy policies online? An eye-tracking experiment. *Computers in Human Behavior*, **55**, pp. 992–1000.

Thomson, M.E. & von Solms, R. (1998). Information security awareness: educating your users effectively. *Information Management & Computer Security*, **6**(4), pp. 167-173

Tötterman, A. & Widén-Wulff, G. (2007). What a social capital perspective can bring to the understanding of information sharing in a university context. *Information Research*, **12**(4).

Urbach, N., Ahlemann, F. (2010). Structural equation modelling in information systems research using partial least squares. *Journal of Information Technology Theory and Application*, **11**(2), pp. 5-40.

Voigt, P., von dem Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR) A Practical Guide.* Cham: Springer International Publishing AG.

Walters, P. (2012). *The Risks of Using Portable Devices*. Carnegie Mellon University.

Wang, Y., Wang, Y., Lin, H., Tang, T. (2003). Determinants of User Acceptance of Internet Banking: an Empirical Study. *International Journal of Service Industry Management,* **14**(5), pp. 501-519.

Wong, K.K. (2013). Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS. *Marketing Bulletin*, **24**(1), pp. 1-32.

Xu, H., Teo, H., Tan, B.C.Y., Agarwal, R. (2010). The Role of Push–Pull Technology in Privacy Calculus: The Case of Location-Based Services. *Journal of Management Information Systems*, **26**(3), pp. 135-173.

APPENDICES

1. Constructs and coded variables used in the research model

Construct	Code	Rating questions	Adopted from
Data Privacy Behaviour	BEH1	Before reading an email, I will first check if the subject and the sender make sense	LaRose et al., 2008; Rogers, 2008
	BEH2	Before opening an email attachment, I will first check if the filename of the attachment makes sense	LaRose et al., 2008; Rogers, 2008
Perceived susceptibility	PSS1	Sharing my personal information with a service provider causes risks to my safety	Heirman et al., 2013
	PSS2	Sharing my location with a service provider causes risks to my safety	Dwivedi et al., 2010; Xu et al., 2010
	PSS3	I am concerned sharing my information on the internet as it might be used for purposes other than I intended	Dinev et al., 2007
	PSS4	I am concerned that my information is used without my consent	Malhotra et al., 2004
Perceived benefits	PB1	Sharing my personal information with a service provider offers me better selection of services	Heirman et al., 2013
	PB2	Sharing my location with service provider offers me better selection of services	Dwivedi et al., 2010; Xu et al., 2010

	PB3	Keeping my devices updated protects my safety	Larose et al., 2008
	PB4	Having a virus software protects my safety	Larose et al., 2008; Lee et al., 2008
Perceived Barriers	PBR1	Practicing care when reading i.e. website's or application's terms and conditions is inconvenient.	Ng et al., 2009
	PBR2	Practicing care when reading i.e. website's or application's terms and conditions is time consuming.	Ng et al., 2009
	PBR3	Practicing care when reading i.e. website's or application's terms and conditions would require considerable investment of effort other than time.	Ng et al., 2009
	PBR4	Practicing care when reading i.e. website's or application's terms and conditions would require starting a new habit, which is difficult.	Ng et al., 2009
Self-efficacy	SE1	I feel confident using social media	Gross et al., 2005
	SE2	I feel confident sharing my location when asked by a service provider	Dwivedi et al., 2010; Xu et al., 2010
	SE3	I understand terms and conditions before clicking "accept"	Steinfeld, 2015

	SE4	I am confident of recognizing a suspicious email or link.	Ng et al., 2009
Perceived seriousness	PSE1	My personal information ending up in wrong hands is a serious problem for me.	Ng et al., 2009
	PSE2	My location ending up in wrong hands is a serious problem for me	Ng et al., 2009
	PSE3	If my computer or mobile is infected by a virus or gets stolen, my safety can be at risk	Ng et al., 2009
Cues to action	CA1	Been afflicted by virus or having my device stolen has caused concerns for my safety	Dodel et al., 2016
	CA2	News and media effect on the way I share my personal information	Glanz et al., 2010; Aytes et al., 2003
	CA3	News and media effect on the way I share my location	Glanz et al., 2010; Aytes et al., 2003
	CA4	My social relations (friends, family, colleagues) effect on the way I share my personal information	Jayanti et al., 1998
	CA5	My social relations (friends, family, colleagues) effect on the way I share my location	Jayanti et al., 1998

2. *The questionnaire*

Questionnaire design

1. Please state your gender
   Male
   Female
   Other

2. How old are you?
   18-25
   26-35
   36-45
   46-55
   56-65
   66-above

3. What is the highest degree or level of school you have completed? If currently enrolled, highest degree received.
   Higher secondary
   Bachelor
   Master
   Ph.D
   Other

4. On average, how much time you spend with personal computer, smart phone or other devices connected to the internet daily?

   Less than an hour
   1-3 hours
   4-6 hours
   7-9 hours
   More than 10 hours
   Not at all

5. How would you rate yourself in terms of familiarity with security practices concerning your personal data? (very familiar/not at all familiar)

Statement	Strongly disagree	Disagree	Slightly disagree	Neutral	Slightly agree	Agree	Strongly agree
Data Privacy Behaviour							
Before reading an email, I will first check if the subject and the sender make sense	1	2	3	4	5	6	7
Before opening an email attachment, I will first check if the filename of the attachment makes sense	1	2	3	4	5	6	7
Perceived susceptibility							
Sharing my personal information with a service provider	1	2	3	4	5	6	7

causes risks to my safety							
Sharing my location with a service provider causes risks to my safety	1	2	3	4	5	6	7
I am concerned sharing my information on the internet as it might be used for purposes other than I intended	1	2	3	4	5	6	7
I am concerned that my information is used without my approval	1	2	3	4	5	6	7
Perceived benefits							
Sharing my personal information with a service provider offers	1	2	3	4	5	6	7

me better selection of services							
Sharing my location with service provider offers me better selection of services	1	2	3	4	5	6	7
Keeping my devices updated protects my safety	1	2	3	4	5	6	7
Having a virus software protects my safety	1	2	3	4	5	6	7
Perceived Barriers							
Practicing care when reading i.e. website's or application's terms and conditions is inconvenient.	1	2	3	4	5	6	7

Practicing care when reading i.e. website's or application's terms and conditions is time consuming.	1	2	3	4	5	6	7
Practicing care when reading i.e. website's or application's terms and conditions would require considerable investment of effort other than time.	1	2	3	4	5	6	7
Practicing care when reading i.e. website's or application's terms and conditions would require starting a new	1	2	3	4	5	6	7

habit, which is difficult.							
Self-efficacy							
I feel confident using social media	1	2	3	4	5	6	7
I feel confident sharing my location when asked by a service provider	1	2	3	4	5	6	7
I understand terms and conditions before clicking “accept”	1	2	3	4	5	6	7
I am confident of recognizing a suspicious email or link.	1	2	3	4	5	6	7
Perceived seriousness							
My personal information ending up in wrong hands is a serious	1	2	3	4	5	6	7

problem for me.							
My location ending up in wrong hands is a serious problem for me	1	2	3	4	5	6	7
If my computer or mobile is infected by a virus or gets stolen, my safety can be at risk	1	2	3	4	5	6	7
Cues to action							
Been afflicted by virus or having my device stolen has caused concerns for my safety	1	2	3	4	5	6	7
News and media effect on the way I share my personal information	1	2	3	4	5	6	7
News and media effect on	1	2	3	4	5	6	7

the way I share my location							
My social relations (friends, family, colleagues) effect on the way I share my personal information	1	2	3	4	5	6	7
My social relations (friends, family, colleagues) effect on the way I share my location	1	2	3	4	5	6	7

3. Cross Loadings Table

	Data Privacy Behaviour	Cues to Action	Perceived Benefits	Perceived Barriers	Perceived Seriousness	Perceived susceptibility	Self-efficacy
BEH1	**0.967**	0.172	0.077	0.077	0.299	0.423	0.086-
BEH2	**0.970**	0.180	0.156	0.107	0.316	0.443	0.156-
CA1	0.070	**0.494**	0.053	0.042	0.311	0.146	0.003
CA2	0.217	**0.832**	0.156	0.020	0.239	0.235	0.175-
CA3	0.185	**0.810**	0.128	0.031-	0.252	0.197	0.178-
CA4	0.079	**0.698**	0.100	0.007	0.216	0.019	0.025-
CA5	0.033	**0.679**	0.034	0.063	0.175	0.090-	0.108
~~PB1~~	0.023-	0.030	**0.330**	0.123	0.086	0.121-	0.098
~~PB2~~	0.040	0.013	**0.417**	0.106	0.042	0.050-	0.092
PB3	0.114	0.171	**0.793**	0.246	0.176	0.225	0.111-
PB4	0.103	0.083	**0.814**	0.150	0.232	0.202	0.078-
PBR1	0.089	0.025-	0.186	**0.870**	0.132	0.171	0.148-
PBR2	0.028	0.062	0.316	**0.869**	0.179	0.129	0.207-
PBR3	0.136	0.012	0.069	**0.706**	0.055	0.193	0.069-
PBR4	0.056	0.082	0.221	**0.704**	0.020-	0.052	0.005-
PSE1	0.315	0.283	0.180	0.055	**0.876**	0.462	0.322-
PSE2	0.270	0.287	0.058	0.085	**0.867**	0.466	0.334-
PSE3	0.191	0.288	0.374	0.237	**0.708**	0.370	0.253-
PSS1	0.445	0.154	0.165	0.204	0.375	**0.838**	0.114-
PSS2	0.381	0.261	0.148	0.112	0.469	**0.827**	0.285-
PSS3	0.350	0.094	0.212	0.203	0.482	**0.820**	0.239-
PSS4	0.289	0.071	0.162	0.110	0.415	**0.814**	0.263-
~~SE1~~	0.075-	0.071-	0.146	0.117	0.021	0.156-	**0.207**
SE2	0.154-	0.109-	0.012	0.106-	0.331-	0.314-	**0.932**
~~SE3~~	0.042-	0.010-	0.035-	0.247-	0.000-	0.089-	**0.111**
~~SE4~~	0.056-	0.000-	0.236	0.192	0.169	0.067-	**0.340-**