On The Adoption of End-User IT Security Measures

Der Fakultät für Elektrotechnik und Informatik
der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

von

Dipl. Inf. Marian Harbach
geboren am 27.04.1985
in Bad Schwalbach

2014

Referent: Prof. Dr. Matthew Smith
Koreferent: Prof. Dr. Michael Rohs
Tag der Promotion: 31.10.2014

Foreword

The work that led up to this thesis would not have been possible without the support of my advisor, colleagues, friends and my family. I am very grateful for the tireless support I received from all of these people, be it through research collaborations, listening to me complaining, giving advice for everything that happened in the past four years or just for being there. This document would not exist without you.

First and foremost I have to thank my advisor Matthew Smith for giving me the chance to work with him and departing on an exciting research journey towards a new field of research. I am also indebted to Michael Rohs, for agreeing to be my secondary examiner. The head of the Distributed Systems institute, Gabriele von Voigt, also always had an open ear for my worries and an open pocket for when I needed money. Similarly, all my colleagues at the Distributed Systems and Security Group, Michael Brenner, Sascha Fahl, Benjamin Henne, Thomas Muders, Christian Szongott, Henning Perl, and Jan Wiebelitz (in no particular order), made for a very pleasant and inspiring working environment. Furthermore, I am thankful for the fruitful cooperation with Alexander De Luca and Emanuel von Zezschwitz at LMU Munich, who shared their experience with me. I also enjoyed collaborating on several projects with my student assistants, Polina Yakovleva, Nils Sonemann, Susanne Weber, and Markus Hettig.

Second, but no less important, my friends helped me to take my mind off of research from time to time and listened to my concerns and complaints. I am very grateful for that. I want to thank Dominik Bilitewski for being a great friend, coffee advisor, and slacklining buddy, Eva Kwoll, Yvonne Niemeyer, Alexandra Wienert, Moritz Herzberg, Christian Neeb, Tim Kaschmieder and Gunther Sander for being an important part of my life and affording me a different perspective on things, as well as Mario Nitschke, Lena Reese, Felix Faust, Christian Bischoff, and Stefan Töhberg for great slacklining and climbing sessions.

Lastly, but most importantly, my parents, Conny and Eckhard Harbach, are the ones who made all this possible from the beginning, for which I will be eternally grateful. And finally Lenka, the love of my life and my wife, without whom I would have gone crazy a long time ago. Your endless and unconditional support and love means everything to me. I dedicate this to you and our daughter.

Hannover, June 24, 2014

Summary

IT security research was centered on developing new technologies for a long time. Only in the past decade have researchers slowly begun to consider human factors when developing and deploying security measures. However, past work has often focused on the usability of the interaction itself. Examples include: how can a password meter become more usable, does changing the text of a warning message increase user understanding, or how do users compose their passwords. A central aspect of IT security measures has however received little attention within the usable security community: why users do or do not adopt a given security measure and which factors influence such a decision in everyday use of technology.

In this thesis, I therefore present a practical view on this important topic based on five studies. First, I investigate users' views on a novel and widely available authentication technology. The findings show that users did not show much desire to use this measure that is actually capable of increasing their IT security while it was already available to them. I present an analysis of two focus groups that show what made this particular technology undesirable for users from a conceptual point of view. Second, I present results on the evaluation of measures to protect the confidentiality of social network communications. The results complement the insights from the first study by showing that small changes in the user interface layout and workflow considerably impact adoption intention.

The third study looks at environmental factors for the adoption of the lock screen in Android smartphones. This everyday security measure is one of the most frequently used security systems and I find that users exhibit a very diverse set of reasons and threats for choosing to use this protection measure or not. The physical measures participants took in addition to the lock screen itself had a great influence on participants' views. The study also finds that the adoption of this measure would increase if protection could be more adaptive to usage context. In the fourth study, the perceived threats and consequences for a more traditional use of the Internet with a desktop browser are investigated. I derive a need for the perception of a graspable benefit in end-users from existing theories on security compliance. In particular, this means that users need to be aware of a threat that can be avoided using a security measure. If no such threat is perceived, motivation for the adoption of this measure will be lower. The results show that which threats users consider relevant when using the Internet is considerably different from the threats security experts consider for their recommendations of security measures to users. This finding can explain the indifference that is repeatedly observed with respect to additional security measures. I also find that the usage context, i. e. which task a user is currently completing, influences which kinds of threats are perceived.

Last, I present the design and evaluation of an improved security measure that is tailored towards users' needs. For this study, the Android permission display was modified to improve the communication of risks and consequences using personalized examples. I find that users stop ignoring this measure and pay more attention to the displayed information in the modified prototype.

Overall, this thesis provides important insights into the human factors that influence the adoption of IT security measures. The presented research informs future design, development, and deployment of user-facing IT security measures.

Keywords: Usable Security, Adoption, Human Factors.

Zusammenfassung

Der Fokus der IT-Sicherheitsforschung lag lange vornehmlich auf der Entwicklung neuer Technologien. Erst im letzten Jahrzehnt haben Forscher erkannt, dass menschliche Faktoren eine wichtige Rolle für die Entwicklung und vor allem die Ausrollung von Sicherheitsmaßnahmen spielen. Der wachsende Bereich der benutzbaren IT-Sicherheit (Usable Security) beschäftigt sich seitdem mit der Benutzbarkeit der Interaktion zwischen Benutzer und Schutzmaßnahme. Es wurde beispielsweise untersucht wie Passwort-Meter effektiver gestaltet werden können, ob das Ändern des Textes in einer Warnungsmeldung die Verständlichkeit seitens der Endanwender erhöht oder wie Benutzer Passwörter erstellen. Eine zentrale Fragestellung in Bezug auf IT-Sicherheitsmaßnahmen wurde jedoch in der Usable Security Forschungsgemeinschaft vernachlässigt: Warum verwenden Endanwender bestimmte Sicherheitsmaßnahmen (nicht) und welche Faktoren beeinflussen solche Entscheidungen in der alltäglichen Verwendung von Informationstechnologie?

Im Rahmen dieser Arbeit präsentiere ich daher Ergebnisse von fünf Studien, die anhand von Case Studies sowie generischen Untersuchungen und der Evaluation eines Prototyps aufzeigen welche Auswirkungen menschliche Faktoren in der Praxis haben können. Zunächst präsentiere ich eine Analyse der Ansichten von Endanwendern zu einer neuartigen und bereits verfügbaren Authentifizierungstechnololgie. Die Ergebnisse zeigen, dass Endanwender kein Verlangen nach dieser neuartigen Technologie haben obwohl sie ihre Sicherheit verbessern würde und sie ihnen sogar bereits zur Verfügung steht. Zwei Fokusgruppen wurden durchgeführt und zeigen welche konzeptuellen Eigenschaften die Technologie nicht wünschenswert erscheinen lassen. Zweitens stelle ich eine Analyse von Maßnahmen zum Schutz der Vertraulichkeit für Nachrichten in sozialen Netzwerken vor. Die Ergebnisse ergänzen die Erkenntnisse der ersten Studie, indem gezeigt werden kann, dass bereits kleine Änderungen an der Benutzeroberfläche und den notwendigen Schritten zur Verschlüsselung von Nachrichten die Bereitschaft zur langfristigen Anwendung dieser Schutzmaßnahme beeinflussen.

In der dritten Studie wird der Einfluss von Umgebungsfaktoren und eigenem Verhalten auf die Verwendung des Sperrbildschirms auf Android Smartphones untersucht. Diese Sicherheitsmaßnahme für Mobilgeräte wird von vielen Nutzern täglich mehrfach verwendet. Die Untersuchung zeigt, dass Anwender eine Vielzahl von mehr oder weniger plausiblen Gründen, Risiken und Gefahren dafür anführen, warum die Maßnahme verwendet wird oder nicht. Die von den Teilnehmern der Studie zusätzlich zum Sperrbildschirm getroffenen physischen Maßnahmen zeigen einen starken Einfluss auf die Ansichten zur Notwendigkeit eines solchen Schutzes. Die Ergebnisse der Studie zeigen auch, dass die alltägliche Verwendung von Sperrbildschirmen erhöht

würde, wenn der Schutz besser auf verschiedene Situationen und Benutzungskontexte angepasst wäre. In der vierten Studie wird das Gefahren- und Konsequenzbewusstsein für die generelle Verwendung des Internets mit einem Browser beleuchtet. Aus existierenden Theorien wird abgeleitet, dass Benutzer eine Schutzmaßnahme nur verwenden werden wenn es einen greifbaren Vorteil bringt, also insbesondere eine Gefahr abgewendet werden kann. Sind die Anwender sich allerdings keiner relevanten Gefahr bewusst, gibt es auch weniger Motivation die Schutzmaßnahme zum Einsatz zu bringen. Die Studie zeigt, dass sich Endanwender bei der Verwendung des Internets anderer Risiken bewusst sind als IT-Experten annehmen und sich die wahrgenommenen Risiken situativ unterscheiden. Dieser Mangel an Bewusstsein kann die immer wieder beobachtete Gleichgültigkeit bezüglich zur Verfügung stehender Sicherheitsmaßnahmen erklären.

Abschließend präsentiere ich den Entwurf und die Evaluation einer verbesserten Schutzmaßnahme für Endanwender. Für diese Studie wurde die Anzeige der App-Permissions auf Android Smartphones modifiziert, um durch die Verwendung von personalisierten Beispielen Risiken und Konsequenzen aus den Handlungen der Endanwender besser verständlich zu machen. Die Ergebnisse zeigen, dass die gewählte Form der Darstellung zu weniger Vernachlässigung dieser Schutzmaßnahme sowie mehr Aufmerksamkeit für die enthaltenen Informationen führt.

Insgesamt erlaubt die vorliegende Arbeit eine tiefe Einsicht in die praktische Relevanz von menschlichen Faktoren bei der Verwendung von Sicherheitsmaßnahmen. Der Entwurf, die Entwicklung und die Verbreitung von Sicherheitsmaßnahmen für den Endanwender können zukünftig von den vorgestellten Ergebnissen profitieren.

Schlagwörter: Benutzbare IT Sicherheit, Akzeptanz, Menschlische Faktoren.

Contents

1 Introduction

> *The most exciting phrase to hear in science, the one that heralds the most discoveries, is not "Eureka!" (I found it!) but "That's funny... ".*
>
> — *Isaac Asimov*

IT security research is perpetually creating novel measures that aim to increase end-user security. However, these measures can only work if such users choose to adopt them. Adopting IT security measures, for the purposes of this thesis, means choosing to use a novel security mechanism, such as a password manager, a chipTAN device, or a fingerprint reader, choosing to comply with security advice, such as creating stronger passwords, not writing down a PIN, or leaving an unlocked computer unattended, or choosing to carefully consider a security decision forced upon the user, such as whether or not to continue to an insecure website, to install an app that requests a lot of permissions, or to allow an unknown program to receive network connections. The reasons for choosing one way or the other and how to incorporate them into improved security measures motivated the research behind this thesis.

To date, there are several theories on why or why not people that use modern information technology choose to adopt a security measure. However, there also clearly still is a lack of adoption of IT security measures, as users still rely on passwords to secure their data, ignore security advice in browsers, and install whatever app they want to use on their smartphone. When helpful security measures are available, but no one wants to use them in daily life, I would say "that's funny... ". However, while the quote at the beginning of this chapter holds true for many problems (and solutions) in science, this particular instance of something funny has only recently become of interest for many researchers in the information security community. Especially, little research has gone beyond the theory-driven work that exists in Information Science and reported on the decision processes pertaining to applying particular measures in everyday life.

Generally, a very large amount of work has been invested in finding better security practices and technologies for decades. With the increasing demand for protection in a world where more and more aspects of our daily lives are supported by some form of IT system, security mechanisms protect an increasing amount of important information. Accordingly, novel and more secure protection measures have been created by researchers around the world. For example, asymmetric cryptography allows for confidential and authenticated transmission of information without the need for exchanging a secret beforehand. Similarly, quantum entanglement can

be used to catch eavesdroppers on communications channels. In the wake of such amazing discoveries, an important aspect has been neglected by large parts of the IT security community: the end-users and their needs. Only in the last 15 years has a community been established that considers how to support the user in the application of IT security in their everyday dealings with IT in general and the Internet in particular. Often, such research however still focuses on applying IT in organizational contexts. Helping end-users at home to protect their digital assets is therefore a very recent trend in human factors and security research, even though a growing amount of services address the needs of this particular audience.

In 2014, almost every single end-user shopping online, using email, playing games, or posting in social networks basically still uses security technology from decades ago. PINs and passwords were introduced with the first automatic teller machines and multi-user mainframes in the 1960s. However, the amount of passwords and PINs a single person has to manage and possibly remember is steadily increasing with every new service one enrolls for online. In 2007, Florencio and Herley [64] found that an average user had 6.5 different passwords and types them about 8 times per day on some of 25 accounts. In the past 7 years, this figure has likely increased significantly. Yet, there are better technologies available to remedy this situations. Digital certificates were introduced, password managers are available, and even the German national ID card was proposed to overcome this burden. Similarly, message encryption using PGP or S/MIME has been available for more than 10 years, yet it has not been widely adopted at all. Also, with the increasing use of IT systems and their existing security technology, users often need to apply this security technology correctly in order for it to be effective. If users simply ignore SSL warnings and app permission screens, these security measures cannot have their desired and important effect. Yet, research has found that this is precisely the case [60, 148].

Meanwhile, the dangers arising from cybercriminals and other online threats are increasing. More and more password databases get stolen, invasions of privacy are commonplace, and botnets comprise millions of infected end user machines. A 2013 PWC survey finds that "the cybercrime threat environment has become increasingly pervasive and hostile"[1]. In this light, it is also surprising that there appears to be little desire for better protection measures from end-users.

But there is also some hope, as researchers begin to tackle this challenge. In 2001, Sasse et al. [140] were the first to argue that security is almost never a user's primary task and that the community needs to rethink security measures. Previously, Whitten and Tygar [161] had already conducted the first and most-cited usability study for a security measure, showing why users were unable to apply PGP 5.0. In the following years, the Usable Security community began to form and in 2008, Cranor and Garfinkel published the first book on this growing topic [35]. Ever since, the awareness of usability issues of security and privacy technology has increased and many positive developments can be observed. Password meters, for example, at least support the user in creating stronger passwords that make account compro-

[1] `http://www.pwc.com/us/en/increasing-it-effectiveness/publications/us-state-of-cybercrime.jhtml` – last access 18.3.2014

mise less likely. Similarly, the number of User Account Control (UAC) prompts was limited to important decisions in Windows 7 and alternative forms of authentication are available on current smartphones, easing the task of unlocking one's smartphone by providing graphical patterns to be drawn or biometric sensors for fingerprints.

However, usability issues and a certain ignorance towards human factors are still present in many IT security measures, as it appears that these aspects are as much a secondary concern for developers and system architects as it is for end-users. Even security measures that have been proven to reduce the burden on users are still not being widely deployed.

1.1 Contributions

In this thesis, I extend the state of the art of human factors in security measure adoption in two ways: I present an exploration of the human factors that influence the adoption of IT security measures for end-users at home and in their daily lives. My analyses increase the understanding of how such users reason about security measures. Additionally, I introduce a novel form of risk communication that overcomes some of the adoption problems that currently exist.

The exploration is based on two lines of work: human factors in general as well as risk perception and communication in particular. First, I describe human factors influencing adoption of new measures in general based on a case study of a real-world and large-scale deployment of a new end-user security technology. I provide new insights into which factors play a central role when a new security system is deployed and which pitfalls should be avoided to find adoption. I also show how existing models for technology adoption match a real-world deployment and provide additional detail to the previously postulated factors.

I also discuss how software design decisions influence how well a measure is received. The results of a laboratory experiment will show how a seemingly minor change in a message confidentiality system drastically reduces its usability and therefore adoption intention. Also, I will show how the design of the user interface has an influence on the trust users will have in a security system. The lesson to be learned from this research is that great care should be taken when choosing cryptographic primitives for end-user security as they can have a major influence on the overall usability and therefore adoption of the system.

In the second part of the exploration, the role of risk perception is investigated. Based on the insights gained from related work and the two chapters on human factors, I investigate the current state of risk perception in users. It is central to a security measure adoption decision how users evaluate the risk of an activity as well as how well they are able cope with it. Only when users see a relevant risk, will they consider adopting a measure to protect against this risk.

To this end, I first present an investigation of risk perception for smartphone lock screens. A survey and a field study show which risks users protect themselves against

using the lock screen or other additional, non-technical measures. The results show that users can carefully consider the risks and possible consequences arising from an activity and choose their desired level of protection accordingly. Additionally, I demonstrate that security measures are part of an ecosystem that includes other ways to increase the security of information. I also find additional evidence arguing for the importance of usability in security measures, as some users spent several hours unlocking their phone over a period of four weeks, even though they only spent a few seconds on each individual interaction with the security measure and infrequently accessed information that they considered sensitive.

Second, I conducted a survey of general risk awareness on Internet users. Existing work in this space always enumerated risks before asking users to rate these. Without prompting for specific scenarios, the survey finds that perceived risks are diverse and dependent on context, are of a non-technical, more general nature, potential consequences of taking this risk are often described in an impersonal way and address a limited set of assets. Furthermore, own negligence is not a source of risk in the users' minds. The survey also shows which risks users are aware of in certain usage scenarios, which can enable developers to address precisely these risks and hence create security measures that users will want to adopt.

Finally, I present a novel user interface that incorporates the results of my previous work to facilitate risk communication using personal examples for app installation. The approach has proven to be effective, increasing adoption of this security measure significantly. This work shows that my research has yielded valuable information for the designers and developers of security measures and also how important it is that users can personally relate to the risks as well as the possible consequences arising from the actions they take.

While I present a successfully evaluated method that increases adoption of a security measure, I must also warn the reader that there does not appear to be a simple solution to the problem of making users adopt security measures more easily. Rather, the non-deterministic nature of the human mind makes security technology adoption a difficult field to navigate. Nevertheless, the insights described in this thesis provide a valuable basis that can inform future developments and deployments of IT security measures.

1.2 About this Thesis

This thesis follows an analytic instead of a constructive approach, as I aim to describe and understand how IT security measures are appraised by end-users. In one chapter, I additionally introduce and evaluate a newly constructed measure that aims to overcome users' disinterest. I chose to follow this path as I felt that gaining an understanding was a more practical and helpful step towards a solution to the lack of adoption for many current measures. A theory-driven, constructive approach to modeling user behavior is also considered "first-wave" HCI and too rigid to capture the diversity of users and usage scenarios that usually occurs for a system [16]. A

more second- and third-wave HCI approach – focusing on user experiences, context, and the impact on culture as well as everyday live – appeared more suitable for this problem, as less rigid methods embrace these differences of a measure's audience and allows for a more practical analysis. I hope my work will prove to be an in-depth basis for other researchers interested in this topic as well as designers and developers attempting to improve on the status quo.

In the three and a half years that have led up to writing this thesis, I have published several papers as first or second author with the invaluable help of many colleagues and friends. Most papers belong in the usable security and privacy domain and also mostly address adoption issues of security measures. For this thesis, I discuss the findings presented in five of these papers that all make a common point: We need to understand how users reason about security measures and make an effort to let users want to have these measures, as forcing them to use particular technologies has been shown to fail. If we succeed to address the actual needs of users, security measures can be more effective. Some of the chapters in this thesis are based on previously published papers and are thus marked with a disclaimer that explains my personal contribution to each of those works. I am, however, greatly indebted to my co-authors for helping me to pursue these ideas. Without them, a large part of my research would not have come to fruition.

The remainder of the thesis is structured as follows. First, I will introduce some background to my work, including the state of the art in end-user IT security as well as in usable security and privacy research, previous work on the adoption of technology, and insights from risk communication and risk perception research. After that, I present a case study of a large scale security technology deployment, namely the new German national identity card and its electronic identity functionality. The results of three focus groups show how diverse the factors that govern the adoption decisions for security measures are and how design decisions of the provider influence the users' views. In Chapter 4, I delve deeper into the details of a message encryption workflow and show how those influence usability and therefore adoption intention of a mechanism. Chapter 5 presents the results of a survey and field study about how users interact with their smartphone's lock screen, why they do or do not use it and how it impacts the mobile experience. Chapter 6 afterwards provides a broader view on a similar topic, namely which risks are users aware of in general when using the Internet for a variety of tasks. I show how a lack of consequences users can relate to may be able to explain the apparent apathy towards many online risks. Finally, Chapter 7 describes the evaluation of an alternative display of smartphone app permissions that made users pay more attention to the contained information and also convinced them to make more security-aware installation decisions. In the last chapter, the findings of my research are summarized before an outlook on future work on the topic at hand is provided.

2 Background

As I endeavor to explore what brings users to adopt certain security measures, some information on the background of the field under investigation is called for. In this chapter, I will hence introduce existing end-user security measures, the relatively new field of usable security and privacy, as well as existing work on the adoption of security measures and risk perception.

Information Technology (IT) Security can be considered a modern form of general Information Security, which is often said to go as far back as the infamous Caesar cipher from 50 B.C.[1] Up until well through the 20th century, information security was mostly a governmental and military application used to protect national secrets against enemies. Possibly the first non-military application of information security was the login password necessary for time-sharing computers, as introduced with the MIT's Compatible Time Sharing System (CTSS) in 1961.[2]

For home users, the first contact with IT security probably came with the first home computer that allowed to lock the screen or user accounts with a password. Ever since the Internet and IT in general permeated many households, IT security has become ubiquitous. Passwords for user accounts on PCs and websites, PINs for mobile phones, TANs for online banking, and access control for Facebook profiles are commonly used by millions or even billions of Internet and computer users.

While there obviously is more to IT security than passwords and Facebook privacy settings [3], it is the parts that users actually interact with which are of interest in this thesis. In the remainder of this work, I will hence use the term IT security measures for those measures that end-users actually deal with.

The remainder of this chapter is structured as follows: First, I introduce the most important, currently deployed IT security measures to give the reader an overview of what is and what is not being widely used by home-users. Next, I give a brief history of the field of usable security and privacy research in order to outline the efforts already undertaken by the community to make security technology easier to use. Afterwards, I introduce existing theories on the adoption of new technology in general and security measures in particular. Finally, as the communication and perception of risk will be a central element of the work presented in this thesis, I introduce previous efforts to shed light on the perception and communication of risk in general and with respect to IT security in particular.

[1] `http://en.wikipedia.org/wiki/Information_security` – last access 7.3.14
[2] `http://en.wikipedia.org/wiki/Password#History_of_passwords` – last access 7.3.14

2.1 End-User IT Security

On the most basic level, IT security can be defined as achieving a set of *goals* or *properties*. ISO/IEC norm 27000:2009 provides only three such goals, namely the preservation of *confidentiality*, *integrity* and *availability* of information. These are commonly known as the *CIA Triad*. However, the norm also notes several additional goals that can also be involved, which are also incorporated into the recent proposal of the *Information Assurance & Security (IAS) Octave* [29]. This list additionally includes *privacy*, *authenticity & trustworthiness*, *non-repudiation*, *accountability* and *auditability*.

To achieve these goals, IT security risks and threats need to be analyzed and appropriate countermeasures need to be taken. ISO/IEC norm 27000:2005 lists the following key activities for the management of information security:

- security policy,
- organization of information security,
- asset management,
- human resources security,
- physical and environmental security,
- communications and operations management,
- access control,
- information systems acquisition, development and maintenance,
- information security incident management,
- business continuity management, and
- regulatory compliance.

Of these eleven items, end-users are generally only exposed to access control and security policies as IT security measures. Access control traditionally comprises identification, authentication and authorization. While I assume that the reader is familiar with these terms, it is important to note that it is mainly these three concepts end-users are concerned with in terms of IT security in the modern Internet. Identification shows a system who I am, authentication proves that it is really me and authorization then grants access to certain resources.

On top of access control mechanisms, a security policy defines what it means to be secure for a system, organization or other entity[3]. Similarly, Anderson [3] defines a security policy as a concise statement of the protection properties that a system must have plus a security target, which is "a more detailed description of the protection mechanisms that a specific implementation provides".

So, for example for a social network site like Facebook, a end-user security policy requires a user of the system to have a unique username for identification (his or her

[3] `http://en.wikipedia.org/wiki/Security_policy` – last access 10.3.14

email address) and a password of at least 6 characters for authentication. Which posts, walls and other data I can access is mandated by Facebook's authorization system that also takes each user's privacy settings into account. Another (fictitious) security policy may require users of a system to identify themselves using a digital certificate and authenticate using the corresponding private key (as a part of a secure authentication protocol). Other security policies may for example also require the user to change the password after a certain amount of time or forbid that users write their passwords down.

It is important to note that these are the minimal security measures service providers on the Internet dictate for their users. Beyond these requirements, users can take additional measures to protect themselves from security threats against their data, their online identities or accounts, and the private systems they own and administer, i. e. devices such as PCs, laptops, smartphones, and tablets. The contents of a security policy may also be a criterion for choosing to use a service or not. Common measures end-users take include backups, anti-virus software, firewalls, and operating-system-level access control as well as password managers to handle the large number of online accounts many users have today or message encryption to improve the confidentiality of communications. With respect to such measures, however, there is no security policy per se and users are free to chose how they want to protect themselves.

Accordingly, the use of available IT security measures differs widely between users. For example, the strong set of the PGP web of trust[4] contains less than 55,000 keys[5]. While this figure does not include users of the alternative S/MIME scheme, this means that only 55,000 users take advantage of the free PGP system and can hence exchange secure emails. Meanwhile, the overall number of consumer email accounts is estimated to surpass 3 billion in 2014.[6] Similarly, Florencio and Herley found in 2007 that users share each of their passwords across 3.9 accounts [64]. This fact indicates a certain fatigue with the password system as users struggle to create individual passwords for each online service they use.

The problem that this thesis addresses hence comprises why users do or do not adopt certain IT security measures within the limits of their discretion. In the remainder of this thesis, I will especially focus on alternative means of identification, authentication, authorization, and confidentiality. The following subsections introduce currently used measures in these areas and available (but often not widely adopted) alternatives.

2.1.1 Identification and Authentication

During their everyday use of the Internet, users most often need to identify and authenticate themselves (i. e. log in). While there is no data available on which identifi-

[4] The Web of Trust establishes identification for confidential emails using PGP.

[5] `http://pgp.cs.uu.nl/plot/` – last access 10.3.14

[6] `http://www.radicati.com/wp/wp-content/uploads/2013/04/Email-Statistics-Report-2013-2017-Executive-Summary.pdf` – last access 10.3.14

cation and authentication technology is most wide-spread, Florencio and Herley [65] surveyed 75 websites from different categories in 2010, including high-traffic sites like Google and Facebook, banks, universities, dating sites, and U.S. government institutions. All of them used username and password for identification and authentication if they offered users to register an account. In October 2013, I also checked the Alexa top 300 websites by traffic[7] and found that of the 242 sites that offer account registration, all use username and password exclusively for this purpose.

Another very common authentication mechanism is used in conjunction with debit and credit cards: Personal Identification Numbers (PINs). Identification is achieved by having possession of the card and authentication by entering the corresponding PIN. The PIN was invented by John Shepherd-Barron in 1967, who came up with the idea of a cash machine for Barclay's bank. The machine would dispense money (up to 10 Pounds) in exchange for a special check in combination with a 4-digit personal secret number. The 4-digit standard remains in place until today and was apparently only chosen because Shepherd-Barron's wife was unable to remember more digits.[8]

However, the plastic card plus PIN combination for identification and authentication was not adopted by consumers for home and online banking. The Homebanking Computer Interface (HBCI) standard was presented in 1995 by the German Banking Industry (DK formerly ZKA).[9] It required the use of a card reader and a software implementing the HBCI standard to secure banking transaction cryptograhically. However, HBCI and its successor, FinTS (Financial Transaction Services), were never widely adopted by end-users. A fact that corroborates this is that, in Germany, all consumer banks primarily offer PIN and TAN (transaction number) based online banking while HBCI/FinTS is only a secondary choice if available at all.[10]

The PIN also made another appearance as mechanism of choice to lock mobile phones. SIM cards in mobile phones had to be unlocked with a 4-digit number in the days of feature phones and modern smartphones simply adapted this mechanism to protect the entire information contained on the phone. Up until the iPhone 5S, a PIN or a password ("passcode") was the only available code-based lock screen on Apple's iOS devices.[11] Devices running Google's Android offer entering a PIN or a password, computing a biometric assessment of the user's face or drawing a pattern on a grid as possible unlock mechanisms.[12] As smartphones become increasingly ubiquitous, a growing number of users has to deal with protection mechanisms many times each day. I will address this issue and why or why not such protection mechanisms are applied in Chapter 5.

[7] `http://www.alexa.com/topsites` – last access 14.3.14
[8] `http://news.bbc.co.uk/2/hi/business/6230194.stm` – last access 14.3.14
[9] `http://www.hbci-zka.de/english/` – last access 14.3.14
[10] `https://www.verbraucher-sicher-online.de/artikel/uebersicht-tan-verfahren-ausgewaehlter-banken` – last access 14.3.14
[11] `http://support.apple.com/kb/ht4113` – last access 14.3.14
[12] `https://support.google.com/nexus/answer/2819522?hl=en&ref_topic=3416293` – last access 14.3.14

Academic Research

There are several studies that describe current practices in online authentication [64, 71, 65, 81] and help to understand users' mindsets. Already in 2006, Gaw and Felten [71] published a paper on the password management strategies of 49 undergraduate students. They found that participants had three or less passwords and were reused only twice. However, password reuse increased over time, as more accounts were registered and reusing passwords made password management easier. They also asked about potential attackers and found that participants viewed people close to them as the most able and motivated attackers. "Unaffiliated strangers" were perceived to be least dangerous by a third of their participants.

In 2007, Florencio and Herley [64] published a large-scale study on how users use their passwords for online services. They instrumented almost 550,000 clients within 85 days and found that users have about seven passwords that they reuse across sites, with each of those being used at six different sites on average. They also found that the stronger a password is, the less it is reused. In their sample, users had to authenticate at the 15 most popular sites in their sample (at the time of writing that paper) every 2.8 to 9.6 days on average, as those sites set cookies that expire only after several days. Another interesting result of this study is that many Yahoo users (about 4.3 %) forgot their password over the three month monitoring period, possibly because Yahoo's password composition policy was rather complex.

These two studies already demonstrated that password use is highly diverse and that users take creative measures to make the password system suit their needs, for example by reusing passwords, using slightly modified versions of the same password or writing the passwords down. Also, the considerable differences in the figures the two studies report show that password habits are rapidly evolving with an increasing use of online services.

As mentioned above, in 2010, Florencio and Herley [65] surveyed the password policies of 75 websites and found that from a service provider's view, security demands are not the primary factor. Services where the user has a choice between several providers or where advertising is a source of revenue allow for relatively weak passwords. The authors conclude that "the sites with the most restrictive password policies do not have greater security concerns", but are "simply better insulated from the consequences of poor usability". They also argue that there is little improvement in terms of security in exchange for the users' inconvenience when demanding stronger passwords. The paper shows that it remains unclear to what extent strong password policies actually increase security.

Finally, in 2011, Hayashi and Hong [81] conducted a diary study of password use. On top of the insights gained from the studies described above, they were able to show how users apply their passwords throughout their daily life and at different devices and in different situations. They recruited 20 participants and gave each of them a small diary that they were asked to carry with them at all times during a two week period. Participants recorded password events (typing a password into any device) in their diaries. During the two weeks, participants typed passwords

75 times on average, mainly to log into online services. Participants had a mean of 8.6 accounts and most password events were associated with email/messaging and online communities. About 85 % of password events occurred at home or at work and in 94 % of cases it was a personal computer or work device.

2.1.2 Authorization and Confidentiality

As introduced above, authorization mechanisms grant a user or another subject access to a certain resource. This resource may be documents or pictures, but can also be actions or capabilities like taking a photo, accessing the address book, or determining GPS coordinates. With authorization mechanisms, the user trusts the system to properly enforce the authorization rules configured by the user. For example, if I configure my Facebook profile so that only other users on my friend list can view my posts, I trust Facebook, the owner of the service whom I entrust my data, that it is really only my friends who can then view the posts. Similarly, when I authorize one Android app on my phone to use my GPS position, I trust the Android operating system (and hence Google, its creator) to not provide my GPS positions to other apps as well. With authorization however, it is not impossible for system administrators to access the data I have entrusted to the service. This means that a Facebook administrator can see all my posts as the Android operating system can always use my GPS position as they are privileged and have special access.

While privileged access to hardware functionality by an operating system cannot be prevented using technical measures, privileged access to sensitive data in communication between myself and another user or system on third party systems can be prevented using cryptography. Cryptographic primitives for encryption allow for so-called end-to-end confidentiality, where no system administrator or even actual attacker can gain access to sensitive data, given that the underlying cryptographic primitives, protocols, and their assumptions hold. Cryptographic measures are usually based on secret key material that only allows those who are in possession of the necessary keys to decrypt the data. While cryptographic measures can also provide additional desirable properties like non-repudiation, integrity and authenticity, it is confidentiality that is of particular interest for the topic of this thesis.

Authorization and confidentiality measures are closely related to privacy. If I want to keep certain information private while it is being handled by an IT system, I need to trust the operator or developer of the system to implement proper authorization mechanisms that keep other unauthorized users of the system from accessing my information. On the Internet, many service providers have privacy policies that describe how user information is handled and who can access which parts of the data for which reasons. These policies are a legal measure that does not prevent abuse on a technical level. To overcome this limitation, users need to encrypt their information to achieve confidentiality and keep it away from prying eyes. Hence, authorization and confidentiality measures are also a form of privacy protection and should therefore be of particular interest for end-users.

From the previous paragraphs, it is clear why authorization and confidentiality are relevant considerations for Internet users. A 2013 representative survey by Pew Research[13] found that 86 % of Internet users have already taken measures to specifically protect their privacy, which underlines that there is a desire for privacy in general. Yet, many authorization and confidentiality measures remain unused. The following subsections will in turn introduce some existing measures that are already in use as well as other measures that are not.

Authorization

The section above already mentioned two forms of authorization that end users are very commonly exposed to today, Facebook's privacy settings and app permissions on Android. With more than 1 billion users worldwide, Facebook and its privacy settings are possibly the most commonly used example. However, it has also been repeatedly shown [114, 115] that users struggle to apply these settings to adequately protect their privacy. Furthermore, Johnson et al. [95] found that the standard differentiation between friends and non-friends on Facebook is inadequate for privacy protection.

A similar picture has been painted about the app permission system on the Android mobile OS. Related work has also shown that users struggle to judge the effects of permissions correctly. Additionally, the repeated exposure to the permission screen makes users ignore this authorization measure altogether. The works of Felt et al. [57, 58, 59, 60] have evaluated this problem in detail. The most striking result is that only 17 % of users in their experiment actually paid attention to the permission screen and only 3 % were able to answer comprehension questions about the permissions. Yet, Android permissions can be seen as a last line of defense before an app is able to access the rich set of personal and potentially sensitive data contained on modern smartphones.

In Germany, there is another common form of authorization: transaction authentication numbers or TANs. These mostly 6-digit numbers are used in online banking to authenticate and therefore authorize individual transactions, such as money transfers. TANs can be delivered to the customers in several different forms, including a paper-based list (classical TAN and iTAN) , as an SMS (mTAN) or using a TAN generator that also requires the customer's banking card (chipTAN). Whenever an online banking user wants to complete a transaction, the system will ask the user to authorize this transaction by entering an appropriate TAN. This form of transaction authorization also includes a second authentication step, as the user needs to be in possession of both, the online banking credentials as well as the TAN list, the generator device, or the registered mobile phone to which the mTAN is sent.

It is important to note that users do not have much choice for or against this mechanism. Banks largely mandate the use of one the TAN systems for online

[13] `http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/` – last access 17.3.14

banking and users appear to prefer this mechanism over FinTS, as described above. In 2012, a representative survey of German online banking practices[14] showed that while 76 % of German Internet users use online banking, only 11 % chose to use FinTS. The remaining users rely on a TAN method. The study also showed that only few users (12-20 %) were considering to adopt alternative technology like the new German identity card or FinTS for online banking. The reasons for this choice, however, remain unexplored.

The comparison of the three examples of common end-user authorization mechanisms shows an interesting pattern. There is Facebook, where users desire to protect their private content from unwanted access but find that the mechanism is difficult to use and does not provide adequate protection in some situations. Yet, users still have control over which information they submit to Facebook in the first place. Then there are Android permissions, where fine-grained authorization of particular functionality is available and a choice is forced upon the user each time an app is installed. However, users often choose to ignore this choice and install an app regardless of which permissions it requests, even though these permissions may mean that an app can access some or all of the private information on the phone. Last, there are the German online banking TANs, which are also forced on the user for each transaction but are still readily adopted as well as desired. One critical difference to the authorization on Facebook or Android however is that users can see an immediate threat to a graspable good, namely their money. Also, transactions will not go through without the user entering the correct TAN, making this security mechanism a direct part of their primary goal. This challenge – security being a secondary goal – is the core challenge of the field of usable security research which will be introduced in Section 2.2 below. Meanwhile, the three examples of authorization mechanisms show that there are factors influencing how security measures are viewed by their end-users which are, to the best of my knowledge, not understood in their entirety at the moment. There are theories on technology adoption which try to explain this conundrum using psychological models and I introduce those in Section 2.3 below.

Confidentiality

Protection of information on the Internet in general and confidentiality in particular is usually viewed from two different standpoints: confidentiality of information in transit and confidentiality of information at rest.[15] There also is the rather theoretic possibility of maintaining confidentiality of information during processing using a technique known as homomorphic cryptography. While there are academic approaches that implement such cryptosystems [23], their lack of performance makes them irrelevant in practice at the moment. Protecting information in transit from network-based attackers is a widely used practice and relies on well known proto-

[14] "Online-Banking – Mit Sicherheit!", TNS Infratest, `http://www.initiatived21.de/wp-content/uploads/2013/01/studie_onlinebanking_fiducia_2013.pdf` – last access 17.3.14

[15] `http://www.sans.org/reading-room/analysts-program/encryption-Nov07` – last access 18.3.14

cols like the Secure Socket Layer (SSL), also known as Transport Layer Security (TLS), or other transport encryption such as IPSec, WEP/WPA(2), or additional VPN tunnel protocols. Confidentiality for information at rest is often achieved using full-disk encryption or using application-specific encryption schemes.

Again, there are several examples of confidentiality measures in use which are more or less well adopted by end-users. For example, major email service providers, including Google and Microsoft have migrated access to mailboxes to secure connections only. Many online shopping sites, such as Amazon or Zalando[16], also rely on HTTPS to secure data in transit, even though this is in some cases limited to the login and checkout process. However, while all modern browsers indicate the security of a HTTPS connection, research has shown that such indicators frequently fail to warn users of security risks [144]. Similarly, Fahl et al. [55] have shown that there are security problems in the use of SSL/TLS by Android apps introduced by app developers. These measures, however, are not fully at the users' discretion but are either offered by a service provider or product or not. Users can then only choose to not use the service because of a lack of security.

In contrast, there are confidentiality solutions which users can choose to adopt by themselves. Examples include whole-disk or email encryption and enabling secure protocols for one's home WiFi network. Some surveys on WiFi security settings conducted in 2011[17] and 2012[18] find that between 18 and 39 % of home wireless networks do not use proper encryption. While there are no current numbers on the adoption of full-disk encryption or email encryption using PGP or S/MIME, Boxcryptor, a software to encrypt cloud storage, allegedly had more than one million users at the end of 2013[19]. However, this is only a small fraction of the 200 million users Dropbox announced to have at the end of 2013[20] and the 250 million users Microsoft reported for their SkyDrive service in May 2013[21]. There also are several confidentiality solutions for social network content, which will be described in more detail in Chapter 4.

While there aren't many dependable numbers on the adoption of confidentiality mechanisms, it appears that adoption is slow in this area. Most emails I send and receive every day are unencrypted and the many users of cloud storage services do not take additional measures to protect their information from the service providers or attackers. How confidentiality measures can be made more usable and therefore more attractive for users to adopt will be discussed in Chapter 4 and which role the

[16] Currently a major online shop for fashion and shoes in Germany.

[17] `http://www.hotforsecurity.com/blog/25-percent-of-wireless-networks-are-highly-vulnerable-to-hacking-attacks-wi-fi-security-survey-reveals-1174.html` – last access 18.3.14

[18] `http://www.safewifi.hk/files/WiFi_adoption_and_security_survey_2012.pdf` – last access 18.3.14

[19] `http://www.wiwo.de/erfolg/gruender/wirtschaftswoche-gruenderwettbewerb-2013-grosses-interesse-an-boxcryptor-wegen-nsa-schnueffeleien/9050318.html` – last access 18.3.14

[20] `http://techcrunch.com/2013/11/13/dropbox-hits-200-million-users-and-announces-new-products-for-businesses/` – last access 18.3.14

[21] `http://blog.onedrive.com/over-250m-people-using-skydrive/` – last access 18.3.14

perception of relevant risks and consequences play will be discussed in Chapters 3 and 6.

2.2 Usable Security and Privacy

While the previous section outlined the basic elements of modern end-user IT security relevant for this thesis, this section briefly introduces the research area this thesis is situated in and provides some pointers on the methods that are commonly applied.

Towards the turn of the century, researchers began to realize that end-users are an important part of IT security systems. Usability and human factors, often referred to as human-computer interaction (HCI), had been recognized as a central aspect of computing in general long before. The first edition of the ACM SIGCHI conference on human factors in computing systems (CHI in short) was held in 1981 addressing how users can interact more efficiently with computing systems. Ever since, this conference has become the major venue for investigations of IT systems' human factors. Creating systems that humans can interact with easily and effortlessly is the core challenge that human-centered design and usability research took on in this space. The importance of the topic is, for example, mirrored in the norm ISO/TR 16982:2002. The norm provides guidance for system creators on methods and concepts for creating usable systems and defines usability as:

> "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use".

The definition underlines that usability is dependent on who is trying to achieve which goals in which context. This acknowledges the fact that a programmer in his office will have a different understanding and way to interact with an IT system than a regular user at home. Usability consultant Jakob Nielsen further deconstructs the concept of usability into five quality components[22]:

- **Learnability:** How easy is it for novice users to interact with a system for the first time?
- **Efficiency:** Having learned how to interact, how quickly can tasks be performed?
- **Memorability:** How easy is it to interact with the system after periods of non-use?
- **Errors:** How many errors occur during interaction? How does the system deal with errors? How hard is it for users to recover from errors?
- **Satisfaction:** How pleasant do users find the interaction to be?

[22] `http://www.nngroup.com/articles/usability-101-introduction-to-usability/` – last access 7.4.14

Usability research strives to maximize these components in order to provide a pleasant experience for users. A quick glance at the history of home computers shows that significant overall progress has been achieved: Since the times of the MS-DOS console with its limited graphics, home computer users can now use graphical user interfaces as well as touch-based interaction on smartphones and do not need to edit configuration files to load drivers anymore.

Human factors and usability did, however, not begin to play a central role for IT security systems and the corresponding research until much later. The work that is probably most frequently cited as "the first" investigation of usability of security systems is Whitten and Tygar's paper "Why Johnny Can't Encrypt" [161] from 1999. In a previous tech report [160], the same authors also first associate usability with security issues in IT systems. Based on a case study of email encryption with PGP 5, the authors raised awareness for usability problems that lead users to corrupt the security of their communications. Only one third of their twelve participants was able to send encrypted and signed emails in their 90-minute test. 25 % of the participants accidentally sent confidential information without encryption. Whitten and Tygar found significant problems with the user interface and questioned PGP's analogy between cryptographic and physical keys. They concluded that the interface *"does not come even reasonably close to achieving our usability standard"* and that it *"does not make [exchanging secure email] manageable for average computer users"*. Based on this work, the authors conclude that "human factors are perhaps the greatest current barrier to effective computer security" [160].

Today, the IT security community is increasingly considering usability aspects and the subfield of *Usable Security* has been established. Cranor and Garfinkel published a first book on the topic in 2008 [35]. Furthermore, dedicated research venues, such as the Symposium on Usable Privacy and Security (SOUPS) as well as the Usable Security Workshop (USEC), have since been created and established security and human factors conferences, such as IEEE S&P as well as ACM CCS and CHI, publish an increasing amount of Usable Security papers.

Previously, usable security research was often concerned with aspects of communicating risky situations to users [20, 21, 22, 133], for example when certificate warnings appear in browsers [76, 148]) or to warn users of phishing attempts [14, 46, 52, 69, 70, 165]. Similarly, how users choose and apply passwords [18, 51, 64, 71, 81, 143] or perceive Android app permissions [8, 57, 58, 59, 60] are two additional examples of important problems addressed in usable security work in the past. All these works have in common that they take a step back from the respective technologies and focus on how users interact with them in order to make recommendations for a more secure environment. Often, non-technological factors were the root cause for many IT security problems. How the usable security community usually conducts their research to deduct such findings will be briefly introduced in the following section.

2.2.1 Methods

An important aspect that separates usable security research from traditional IT security research is the choice of methods applied. While researchers have been tackling IT security problems with formal security analyses and mathematical proofs of cryptosystems, usable security researchers rely more on the sociologist's and psychologist's toolbox. Empirical methods, such as surveys, interviews, and focus groups are applied but also experiments conducted in labs, in the field or online. These methods are also commonly used in HCI research. The interested reader is referred to a focused introduction to empirical methods for human-computer interaction problems by Lazar, Feng and Hochheiser [112].

A major trend in HCI research is the use of crowdsourcing services in general and primarily Amazon's Mechanical Turk[23] (MTurk) in particular for HCI research [104] including usable security problems [99]. Willing Internet users, so-called workers, can choose to work on "human intelligence tasks" (HITs) which were originally designed to provide HIT requesters help with tasks computers cannot easily solve, such as categorizing whether or not an image contains any sky or writing short texts about a product. For each completed HIT, the worker is payed a small amount of money – usually between 1 cent and $ 10 – after the requester has approved the result. However, this service has also been increasingly used to present simple online experiments or surveys to a wide population of Internet users in the United States. While the use of MTurk does not ensure the representativeness of data as demographics differ from the general US population [139], it enables researchers to have easy access to a large participant pool, which was often not trivial in the past [99]. It is also important to take appropriate precautions when using MTurk as a research platform, as workers are motivated to finish tasks as quickly as possible to make more money [49].

For the purposes of this thesis, it is important to note that a central challenge in tackling usable security problems is the choice of suitable empirical methods and designs from the large arsenal of available possibilities. As this choice is closely related to the problem at hand, the methods applied in this thesis are discussed within the respective chapters. In the next section, I will delve deeper into existing work that addresses the central problem tackled in this thesis, namely what influences the adoption of IT security measures by end-users and how the risks such measures protect users from are perceived.

2.3 Adoption of Technology and Security Measures

In this section, I will present three streams of work that pursue a goal similar to that presented in this thesis. First, there are models from Information Science and Management based on psychological theories that try to explain why a certain technology is adopted ("accepted" in the terms of this line of research) or not. I will introduce

[23] `http://www.mturk.com` – last access 7.4.14

the underlying psychological models as well as recent models for information technology in general and security technology in particular. It is important to note that the research conducted in this space is often situated in organizational contexts, as managerial aspects often play a role. Afterwards, I present the work surrounding the compliance budget, which tries to explain the behavior of users around security policies in corporate contexts. The authors of these studies argue that people only make rational choices when ignoring organizational security policies. Last, there are several independent reports of research on the adoption and the reasoning about security technology, which I will present in the third part of this section.

2.3.1 Psychological and Information Science Models

There is a considerable amount of research on technology adoption in general from the area of Information Science and Management. This research is often based on psychological models which have been created for different applications but have shown utility for additional purposes. The most recent model for technology adoption that explicitly addresses IT security services at the time of writing has been proposed by Herath et al. [83]. They combine the well-known and widely discussed Technology Acceptance Model (TAM) of Davis et al. for general technology acceptance [38] with Liang and Xue's Technology Threat Avoidance Theory (TTAT, [113]), which relates avoiding security threats to coping behavior and Protection Motivation Theory, which are in turn models from Psychology.

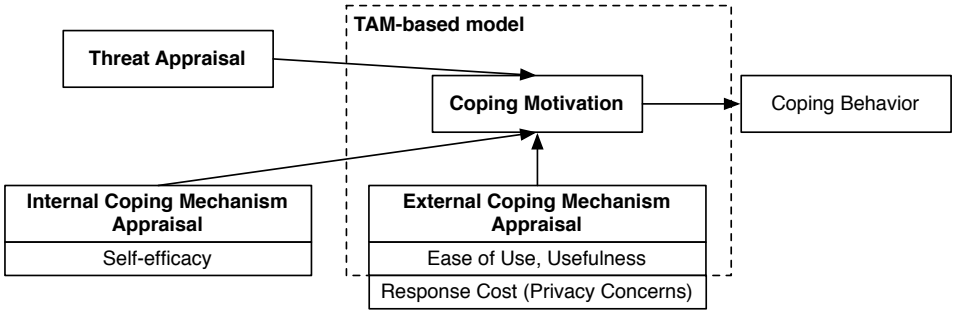


Figure 2.1: The model presented by Herath et al. Image adapted from [83].

In Herath et al.'s model (cf. Figure 2.1), users' motivation to adopt a mechanism depends on three main factors: threat appraisal, internal coping mechanism appraisal and external coping mechanism appraisal. This means that users need to decide to what extent a threat applies to them, evaluate whether or not they can cope with this threat using existing – or internal – mechanisms, and whether or not another – or external – mechanism is suited to cope with this threat. In the model, the appraisal of external mechanisms mainly depends on the two main factors outlined in TAM: Ease of Use (i. e. Usability) and Usefulness (i. e. Response Efficacy). Additionally, Herath et al. posit that response costs, for example concerns about the violation of privacy, play an important role in this process. These three components

then predict coping motivation which will ultimately lead to Coping Behavior, i. e. the adoption of security technology.

The original TAM of Davis et al. [38] mainly consists of the two components mentioned above: Ease of Use and Usefulness. According to the model, these influence each other and are also moderated by so called external variables, which mainly comprise additional information sources. Eventually, an *Attitude Towards Using* as well as a *Behavioral Intention to Using* are formed which then leads to system use. Figure 2.2 visualizes these relationships. TAM itself is based on the Theory of Reasoned Action (TRA) [62], which assumes that "behaviors are best indicated by intentions and those intentions are formed jointly from the constructs of subjective norms and a person's attitude" [31].

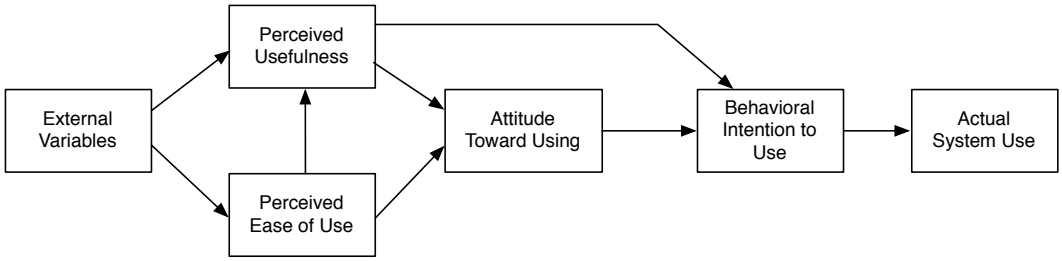


Figure 2.2: The Technology Acceptance Model of Davis et al. [38].

The other basis for Herath et al.'s model is the TTAT by Liang and Xue. In their theory, the authors postulate that the use of security technology is not about adopting a technology but about avoiding a particular threat. This is different from other theories on technology adoption as the authors acknowledge that the users' goal with IT security technology is not to facilitate some task or improve an aspect of computer use, it is about avoiding negative consequences. The important conclusion from this view is that there is more than one path to attain this goal. This is where coping as a mechanism to deal with IT security threats comes into play. As Figure 2.3 shows, a need for action (threat and coping appraisal) will lead to taking precautions, either by emotionally coping with the problem (i. e. accepting the threat and its consequences, dismissing it as irrelevant, etc.) or by taking a problem-focused approach (i. e. actively pursuing measures to make the problem go away). Most importantly, TTAT shows that perceiving a threat to be existent is a necessary precondition for beginning to cope with it and coping can take many forms, including dealing with the problem emotionally or actually adopting a new safeguarding technology.

The threat and coping appraisal part of TTAT is in turn based on Protection Motivation Theory (PMT), which was first described by Rogers in 1975 [137]. It was originally created to explain how appeals to people's fears affect decision making in health-related topics. PMT can be applied to evaluate how effectively recommendations to avoid negative consequences are received. It uses the perceived severity of a threatening event, the perceived probability of the occurrence, or vulnerability, the efficacy of the recommended preventive behavior, and the perceived self-efficacy

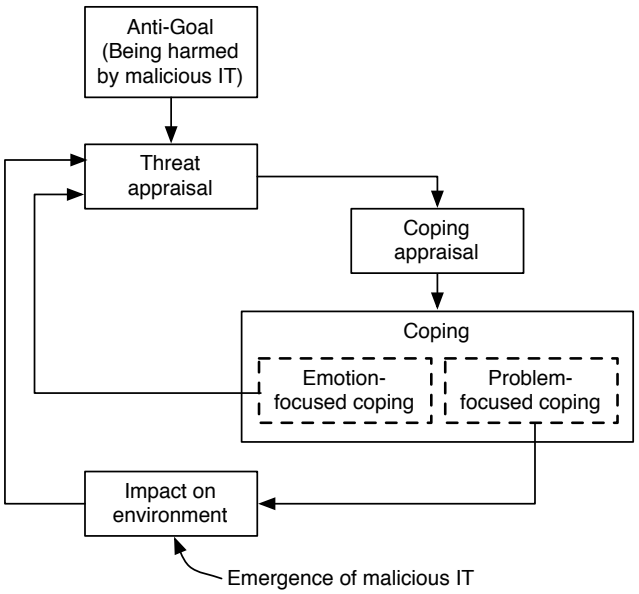


Figure 2.3: The Technology Threat Avoidance Theory (TTAT) of Liang and Xue. Figure adapted from [113].

to deal with this event to explain how motivation to protect arises (cf. Figure 2.4). Claar [31] finds that "the PMT is well suited to study protection behaviors outside the health care field", for which it was originally created, as it "has been applied to over 30 different domains, both inside and outside the health-related contexts". With respect to IT security technology, several studies have applied PMT to show a relationship between threat appraisal, coping appraisal, and motivation to adopt a security technology (e. g. [28, 36, 111, 159, 164]).

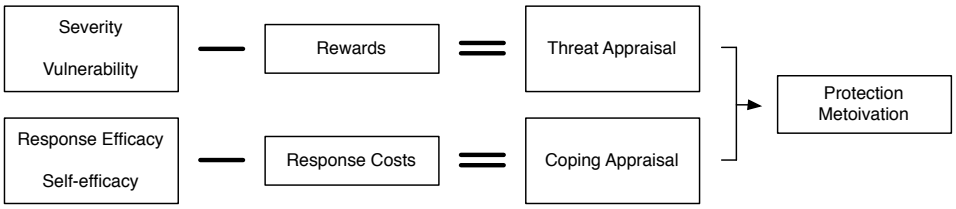


Figure 2.4: The Protection Motivation Theory (PMT). Figure adapted from `http://en.wikipedia.org/wiki/File:Protection_motivation_theory.png` - last access 11.4.14.

An earlier theory that has also been applied to model adoption of IT security measures is the Health Belief Model (HBM) [86]. Its core constructs are very similar to PMT, but with an increased focus on people's beliefs and attitudes. Also, the HBM postulates that cues to action (media reports, personal influences, etc.) are necessary to release a potential for adopting a behavior to reduce a threat. Ng et al. [126] have applied HBM to study computer security behavior and found that the postulated factors were determinants of email security behavior.

The work of Claar [31] gives a more detailed overview of these and additional models. The main difference between the approach to security technology adoption in the works outlined above and the results presented in this thesis is top-down versus bottom-up. Psychological and behavioral models start from a general theory and try to fit observed behavior into these models. Additionally, these models use Likert scale agreement to evaluate the amount of variance in answers that can be explained by the postulated factors. To which extent people's real intentions and future behavior can be extracted from choosing a level of agreement on a numeric scale is at least debatable and not common practice in the usable security community.

The work I present in the remainder of this thesis takes a bottom-up view that may not be comprehensively applicable across different usage scenarios, but it yields detailed insights into particular technologies and also provides a more viable basis to create solutions for the adoption problems in these very areas. My work also shows that the factors postulated in the models can inform the design of usable security systems and that a more detailed understanding of how these factors manifest themselves in real-world deployments is important.

Additionally, all of the models presented above show that users need to be aware of a particular threat in order to arrive at a positive evaluation of a corresponding protection measure. The work I present in Section 6 shows which risks current Internet users are actually aware of and how the lack of awareness I found may be able to explain the apathy towards many security measures.

Beyond models to explain adoption of security technology, Herath and Rao as well as Pahnila et al. also used PMT and similar theories to model the compliance with security policies in organizations in a top-down fashion [84, 128]. To the best of my knowledge, this is the only stream of work in this area that investigates the adoption of security measures beyond concrete products and technologies. Again, the authors find that the postulated factors influence self-reported intentions to comply. However, policy compliance is still indirectly operationalized as agreement to statements like "I am likely to follow organizational security policies". It can be problematic to assess why users do not adopt certain measures by asking hypothetical questions. Usable security research generally focuses on more direct operationalizations by observing actual behavior instead. Nevertheless, the factors postulated in the models presented in this subsection in general can also be found in the results I will present in the remainder of this thesis. In the next subsection, I will introduce previous work from the usable security domain that pursues alternative approaches to explain the adoption of particular measures.

2.3.2 Security Measure Adoption in Usable Security Research

In the usable security domain, why users do or do not adopt a security measure has been mainly investigated based on security policies and security advice. In 2001, Sasse et al. already argued that a lack of compliance occurs because security is not a primary goal for the user [140]. In follow up work, Beautement et al. [7] presented the concept of the Compliance Budget. The theory is based on interviews with 17

regular employees of a big UK company and applies to situations where users have a choice about whether or not to comply with security regulations. The authors argue that this is the case for most companies, as compliance can seldom be constantly monitored. According to the theory, employees are ready to invest a limited amount of effort (their budget) into complying with security policies. They will weigh the cost of complying (usually additional effort) against the benefit of doing so. Whenever the costs exceed the benefits – which is usually the case according to the authors – the difference is deducted from their compliance budget. If the budget is exhausted because they are already complying with other regulations, employees will choose to ignore a new policy.

The authors argue that the best way to improve compliance with security policies is to attempt to tip the scales in favor of the benefits. They propose that security should create minimal extra effort by design, awareness should be raised to increase the perceived benefit of security measures, a security culture should be established in the organization making compliance seem desirable, and monitoring and sanctions should be in place so that avoiding these becomes a benefit of complying.

Similarly, Herley [85] extends this reasoning to regular computer users and argues that these users make a rational choice when ignoring security measures. The author analyzes costs and (potential) benefits of password composition advice, checking visited websites for signs of phishing, and minding certificate warnings in the browser. He finds that costs are greatly exceeding the potential benefits. According to the analysis in this work such a view is created in users because there is no evidence that users who follow security advice fare any better than those who do not. Additionally, Herley argues that many experts do not consider that user effort is not free but comes at a cost in terms of time that can be better spent. Also, many pieces of security advice protect against worst case scenarios which rarely occur in the reality of individual users. As a recommendation, the author also proposes to redesign security advice and measures for regular computer users to create benefits they actually care about. The work I present in Section 6 is another important step in this direction.

Other past work in the usable security domain has also investigated adoption of authentication technology: Jones et al. described users' general perceptions of authentication mechanisms [96]. They find that users have limited knowledge of available alternatives for authentication and showed more positive views when financial applications were to be protected with improved technology. Online shopping applications, in turn, were perceived to be less in need of sophisticated protection technology. These results suggest that users do differentiate usage contexts and are less ready to invest effort depending on the value of the assets to be protected.

Weir et al. compared the usability of two-factor authentication for eBanking applications [158] and found that participants chose the least secure, but most usable method as their preference. The authors speculate that this is because the increase in security does not justify the added complexity of the respective methods, even when the most tangible asset (i. e. money) is involved. They also found that reusing their preferred method a second time made participants perceive greater usability

with respect to the other available methods. The authors conclude that it may be beneficial to offer users a choice of methods in order to make them feel more satisfied with their own choice.

Summary The work presented in this section outlined several views of how users make decisions about security measures. A common theme of all approaches is that there is an appraisal of the threat as well as the proposed protection mechanism. Only if a suitable balance is struck between these two factors will a technology, a policy, or a piece of advice be adopted. As described in the introduction, one way to tip the scales in favor of additional security is to improve user understanding of the risks they can protect themselves against. Research on how to communicate risks to users and user perception of IT security risks is outlined in the next section.

2.4 Risk Communication and Perception for IT Security

In IT security, the communication of risks is usually achieved by displaying warning messages for the user at some point during the use of a software. The work of Bravo-Lillo et al. [20, 22] detailed how warning messages influence user behavior and also presented an improved design that overcomes the limitations of existing approaches [21]. Similarly, the efficacy of certificate warnings in browsers [76, 148], warnings of phishing attempts [52, 165], website security indicators [144], and Windows' User Account Control [124] have been investigated. The results mainly found that users do not heed the warnings and also do not read the text included in such dialogs. Reasons identified for this behavior include [19] too complicated texts [76], inability to understand what the warning is about [48, 124, 134], users' inability to understand what the provided options entail [52], lacking awareness of risks [48] or differing mental models [52, 165], and an underestimation of risks [134].

More recently, research focused on risk communication beyond warnings in the form of education and visual cues in order to communicate IT security risks to users more efficiently and let users make better informed security decisions. For example, work by Blythe et al. and Garg et al. investigated how to improve communication of a phishing risk to elderly users using a video [14, 70]. De Luca et al. [39] used a back-lit keyboard to indicate insecure websites containing forms to be filled. The keys would be red when the site was insecure and hence dangerous for the user to enter sensitive data and they found that security behavior improved. Yet, their prototype also caused some confusion among participants as the visual metaphor of glowing keys was not always immediately clear to the users. Maurer et al. [120] used data-type specific visual indicators immediately next to sensitive HTML form fields, for example when credit card information needed to be entered. This additional attention and awareness immediately next to the information in question made users behave more securely. Lastly, Raja et al. [133] analyzed users mental models of personal firewalls and created visual metaphors for firewall security indicators in order to replace simple green and red lights or textual descriptions. These approaches

allowed the usable security and privacy community to instill a certain amount of awareness for particular security and privacy measures in users. The work I present in Chapter 7 also aims to create more awareness and attention for a security-relevant decision.

Additional research investigated which factors influence the perception of enumerated risks. Huang et al. [89] identified six factors that influence the perceived danger of risk like Denial-of-Service (DoS) attacks, hardware failures, or natural disasters. The identified factors comprised Knowledge, Impact, Severity, Controllability, Possibility, and Awareness. Garg and Camp [69] leveraged the model of Fischhoff et al. [61], which was created to investigate the fear people have of nuclear energy, and applied it to identify the contributing factors for the perception of IT security risks. Fischhoff et al. originally extracted nine dimensions which Garg and Camp adapted to IT security:

- **Voluntary:** Are people subjected to these risks voluntarily?
- **Immediacy:** Is the risk from the threat immediate or does it occur at a later time?
- **Knowledge to exposed:** To what extent are the risks associated with these threats known by the people exposed to them?
- **Knowledge to science:** To what extent are the risks associated with these threats known to experts, like computer scientists?
- **Controllability:** If you are exposed to this threat, to what extent can you control (or mitigate) the risk, by virtue of personal skill or diligence?
- **Newness:** Are these risks new, novel ones or old, familiar ones?
- **Chronic-catastrophic:** Does this risk affect one or multiple systems at a time?
- **Common-Dread:** Is this a risk that people have learned to live with and can think about reasonably calmly, or is it one that people dread on the level of a gut reaction?
- **Severity of Consequences:** How severe do you think the consequences would be if this threat were exploited?

When evaluating to which extent these factors determine the perceived riskiness of 15 common online threats including viruses, phishing, and surveillance, they were able to increase the explanatory power of the model by reducing it to four dimensions:

- **Familiarity:** Knowledge to science and Knowledge to the exposed;
- **Impact:** Severity, Chronic, and Immediacy;
- **Control:** Voluntariness and Controllability; as well as
- **Temporal Impact:** Newness and Common-Dread.

Overall, they find that the less familiar a risk is, the less control the user has over the source of the risk, the higher the impact of the risk's consequences and the newer

and less common the risk is, the worse participants perceive this risk to be. The authors note that the temporal impact contributed the most explanatory power.

There also is existing work on risk perception that analyzed users' mental models [156] or how to incorporate these into security solutions [15] with respect to specific threats, such as phishing, hacking or malware. Dhamija et al. [46] have argued that phishing works because users do not sufficiently understand the technology, and therefore its risks.

2.5 Summary

In the first part, this chapter introduced some background on IT Security in general and the relatively new area of Usable Security in particular. In the second part, I summarized the existing related work around the topics of technology adoption as well as risk communication and perception. The considerable amount of work from several disciplines provides a good basis to investigate specific aspects of security measure adoption further. Many of the results presented above relate to a particular product, technology or scenario and often rely on a particular framework to make a point. For example, technology adoption models are built on existing theories and are refined for a particular technology, such as email authentication in the work of Herath et al. [83]. These models also rely on rating scales completed by a limited population in a survey to predict behavior, which can be problematic as the intention to do something is not necessary equal to actually doing it at a later point in time. Similarly, research on risk perception and awareness has focused on understanding how a set of enumerated risks is perceived by users or how a particular threat can be more effectively communicated. Users' general state of awareness – without asking them how risky they, for example, perceive phishing to be – has not been explored at all. The work I present in the following chapters extends the state of the art as follows:

- I investigate the (lack of) adoption of a real-world, large-scale deployment of a security technology in Chapter 3 that is technically superior to existing measures;
- I present results on how seemingly insignificant changes in a security workflow for message confidentiality influence adoption intention in Chapter 4;
- I analyze security behavior, risk perception, and motivation to use additional security measures during the use of smartphones, combining an online survey with a four-week field study in Chapter 5;
- I discuss the state of general risk awareness in users of the modern Internet, without priming them about particular threats in Chapter 6; and
- I present and evaluate a novel user interface that leverages personal examples to improve risk communication in Chapter 7.

3 Human Factors in a Large-Scale Security Technology Deployment – A Case Study of the German eID Card

This chapter investigates why many users reject a new measure that was intended to conveniently improve their IT security. Among many human factors, the chapter also shows how the ecosystem of a protection measure and social factors influence adoption.

3.1 Motivation

As I argued in the previous section, username and password remains the prevalent mechanism for every-day online authentication. While services and usage patterns evolve, this authentication mechanism remains largely unchanged throughout the history of Information Technology. Users entrust private information to online services and protect that information with their login information. Authentication is therefore a crucial aspect of everyday human-computer interaction as well as the corresponding security mechanisms.

Of Alexa's top 100 websites[1], most sites offer additional features behind a username-and-password-based login or require a login to access the site at all. Users have learned to accept this form of authentication [81] and use creative schemes to tailor the system to their needs [118]. They use separate pseudonyms for different services and choose password strength according to several criteria [64, 71, 81] to maintain appropriate levels of privacy and security.

Several papers (e. g. Bonneau [17], Dhamija et al. [45], Jakobsson and Dusseault [93], as well as Perito et al. [130]) and incidents[2] have shown problems with the current practice and numerous alternative online authentication schemes have been proposed to overcome these [44, 18]. However, none of these schemes has found wide-spread adoption yet. In an extensive survey of proposals for improving online authentication, including password managers, graphical, cognitive, phone-based or biometric schemes, as well as paper and hardware tokens, Bonneau et al. [18] found that none of the 35 mechanisms under investigation came close to the bene-

[1] `http://www.alexa.com/topsites/global` – last access 12.6.14

[2] `http://www.h-online.com/security/news/item/Password-leaks-bigger-than-first-thought-1614516.html`, `http://arstechnica.com/security/2012/08/passwords-under-assault/` – last access 7.3.14

fits username-and-password currently offers to users. While security was generally better with the alternatives, deployability was always worse and with many usability also suffered. While hardware tokens such as smartcards could provide strong security, these mechanisms have high deployment costs for both users as well as providers and additionally require the user to carry an object around.

A trend of the past years could, however, be able to alleviate these major downsides of physical tokens. Many countries are currently in the process of rolling out or have already rolled out national identity cards with a means to access identity information electronically (electronic identity, eID). Acuity Market Intelligence projects that 85 % of identity credentials issued annually will be electronically readable by 2015 and four out of five countries will be issuing eID-capable identity cards[3]. This implies that, soon enough, a considerable portion of Internet users will be carrying a token, capable of electronic authentication, that has already been paid for.

While the USA have not yet announced plans to introduce eID documents, 24 of the 27 countries in the European Union have already deployed or plan to deploy eID cards[4]. The European Commission is actively supporting digital identities and aims to adapt legislation to increase adoption[5]. Since national eID cards are compulsory in many states, they have the potential to be broadly adopted as an electronic means of identification. Additionally, these cards are often habitually carried by their owners and hence, similar to mobile phones, are more or less ubiquitously available. Currently, available solutions in countries such as Estonia or Belgium aim at proving real identities for public services/eGovernment as well as eCommerce and eBanking applications. This focus has however limited their applicability as a general online authentication mechanism.

In this chapter, I analyze the most advanced eID scheme: the nPA (*Neuer Personalausweis*, new personal identity card) launched by the German government in November 2010. This official document is the first to include privacy-preserving features beyond proving one's true identity electronically. While the nPA can reliably verify that a person lives in a certain city or is of certain age without disclosing the actual address or age, it can also generate a provider-specific pseudonym directly on the card. It is therefore the first national ID card technically capable of substituting username and password without any threat to the user's privacy. With the nPA, the user can choose to login without disclosing identifying information or with dependable proof of his true identity. Also, using the nPA would free users of the need to remember a larger number of passwords for the many services they use from day to day.

The aim of this chapter is to analyze which factors inhibit the adoption of the nPA. Given that this technology is part of a large-scale government initiative, it has a certain momentum and therefore potential to gain traction, at least in Germany.

[3] `http://www.acuity-mi.com/GNeID_Report.php` – last access 12.6.14

[4] `http://www.epractice.eu/en/community/eureid/view_resources/Factsheet-on-the-electronic-Identity-at-pan-European-level---May-2012` – last access 7.3.14

[5] `http://www.euractiv.com/infosociety/brussels-wants-identities-eu-cit-news-512833` – last access 7.3.14

Using this technology as a case study, important design decisions, failures, as well as miscommunication can be named that future developments may be able to avoid. The core research contributions of this chapter are as follows:

- I investigate users' perceptions in using a new system for everyday online authentication and identify motivations for, as well as barriers to adoption.
- I present technical and social factors that influence the adoption of alternative authentication mechanisms in general and how privacy-preserving features compare to these other factors.
- I shed light on the needs and perceptions of businesses offering such an alternative authentication mechanism for their online services.
- I provide recommendations that can support the deployment of future authentication systems and also shed light on factors involved in the adoption of security measures in general.

While I do not suggest that the nPA is the best possible solution, I posit that smartcards issued by national governments can provide a viable basis for more security and good usability in online authentication. Germany's efforts to maximize privacy in their eID scheme is an important step towards enabling wide-spread use.

The remainder of this chapter is structured as follows: The next section presents related work of special interest for this chapter. Afterwards, I outline the German eID scheme and show how much care has been taken to create a privacy-preserving solution. I then introduce research questions and the findings of a user study on user perceptions concerning the nPA and also look at factors hindering adoption from the business perspective. Finally, I summarize the findings and suggest an extension to Herath et al.'s [83] model for technology acceptance to explain the technical and social challenges involved in finding wide-spread adoption for a novel, general-use authentication technology.

Disclaimer: The contents of this chapter were previously published and presented at the Privacy Enhancing Technologies Symposium (PETS) in 2013 under the title "On the Acceptance of Privacy-Preserving Authentication Technology: The Curious Case of National Identity Cards" together with Sascha Fahl, Mathias Rieger and Matthew Smith [74]. The co-authors advised me while developing the study design, helped with conducting the study, and provided feedback on the manuscript. At the time of publishing this paper, it was the first to discuss the role of national ID cards in the search for a username-and-password alternative.

3.2 Related Work

On top of the work on technology acceptance models and other investigations of security technology adoption presented in Chapter 2, there is additional related work concerning eID, its application and federation as well as alternatives to the current username-and-password practice.

The challenges for digital identity management have been discussed in general by Dhamija and Dusseault [45] and for eID in particular by Grote et al. [72]: While, according to both these works, the eID functionality of Germany's new ID card has flaws, the scale of the roll-out and the implicit trust in a government-issued document offers an opportunity to investigate the acceptance of a real world deployment above and beyond purely academic proposals. Yet, as I will show, barriers to adoption currently prevail.

Sun et al. [147] ran a comparative user study investigating usability issues of another alternative form of authentication. They found that users need more privacy control, better integration and trust in the involved parties when using OpenID. In contrast to eID, OpenID identity providers need to gain the user's trust. Additionally, identity providers often offer their services as a secondary function and are therefore not per se perceived as an entity providing identity assurance.

Dey and Weis [44] presented an approach to add pseudonymity to OpenID federations and Perito et al. [130] demonstrated that it can be easy to link pseudonymous and public user profiles across services to gain additional information about users. These papers argue for the need of privacy in online authentication, which the German eID scheme is able to offer to a wide audience, hence eliminating one hindrance proposed in other research.

Furthermore, the EU projects PRIME and PrimeLife addressed many facets of digital identity management. In particular, Wästlund et al. [157] analyzed several UI metaphors to communicate privacy-enhancing features in identity management to users. The work presented in this chapter complements the work of Wästlund et al. by investigating the role of privacy-preserving features in the light of other adoption factors in a real-world, large-scale roll-out of a new authentication system.

Historical and sociological aspects of the perception and adoption of electronic identity cards have been investigated by Bennet and Lyon [9]. To increase the utility and applicability of eID, the STORK and STORK 2.0 projects have been continuously working on standards to make the individual eID projects of EU member states compatible and to enable cross-border interoperability. Furthermore, the Open eCard App [90] and two smartphone App projects[6] show how software integration might help to overcome the issue of interoperability and card reader availability.

3.3 The German eID Scheme

The information contained on the new German ID card includes first and surnames, title, address, date of birth and document type.[7] Additionally, the card can provide functionality for residence verification, age verification and pseudonymous identi-

[6] `http://www.cased.de/en/press/archive/29` and `https://play.google.com/store/apps/details?id=de.bos_bremen.android.autent.pinapp`

[7] Currently, the only document type is the ID card itself. However, if eID functionality was integrated into other types of official documents (e.g. driver's licenses), this field would indicate which type of document the user is authenticating with.

fication without disclosing the actual information [67]. It is the first eID-enabled document in the EU to provide such functionality while including privacy and security as major design goals.

Margraf [119] lists the requirements adhered to during the design process of the new card, including encryption of all transmitted data, explicit user consent to all data transmission, authentication of the communication partner, transmission of necessary data only, inability to monitor the card holder's activities, revocation of lost cards, the ability to provide pseudonymous authentication, and the infeasibility of identifying the user and the card through unique properties. These requirements are met using a number of technologies (Card-Verifiable Certificates, Extended Access Control) and organizational processes. The interested reader is referred to the technical specification [26]. In the following, I present the online authentication process using the nPA from a user's perspective and will also briefly outline the requirements for a service provider.

To authenticate to a service, the user needs three things: his ID card with the corresponding PIN, a certified card reader and a certified client software to run the necessary protocols. The card reader costs between 30 Euros for a basic reader without keypad and 160 Euros for a model with display and keypad. This is also one of the major disadvantages of the scheme which I will discuss below. A free reference implementation of the software, called the *AusweisApp*, is currently available from the Ministry of the Interior. Although the need to memorize a PIN may cause problems common to password schemes, users would only need to remember one PIN instead of a larger number of passwords. Users are already used to remembering PINs through the use of banking cards and mobile phones.

When the authentication process begins, the service provider requests authentication using an eID service run by a certified third party. The eID service initiates the cryptographic protocols and establishes a secure channel between the card and the service provider. The user is then presented with information on the service he is authenticating to by the AusweisApp and can also verify which of the information on his nPA will be shared. The user enters the personal PIN code and the client software securely delivers the desired information to the service provider.

In order to use eID functionality, a service provider needs to apply for an authorization certificate at the Federal Office for Authorization Certificates. The provider's need for identifying data will be evaluated by examining her business processes. Service providers get an authorization certificate if they can demonstrate a need for the requested information in compliance with privacy regulations. Only this certificate then enables a service to read exactly the approved data items from the ID card. Currently, 85 providers hold one or more authentication certificates. It is important to note that almost any service provider would be able to get a certificate to access the pseudonymous identification function, which does not disclose any personal information to the service provider but only allows to re-identify a returning user.

In November 2011, after one year, 8 million citizens have already received an nPA[8]. Based on internal statistics, interview partners from public offices (see below) estimated that in September 2012 up to 15 million German citizens were in possession of a new eID-capable personal identity card. These sources also reported that only about 20 to 30 % of card holders have their eID capabilities activated, since eID can be deactivated upon request when the user picks up the card; I discuss this further below.

3.4 Research Questions

In this chapter, I offer new insights by examining the users' perceptions of a large-scale and secure scheme that they already have access to but are reluctant to put into use. Investigating the adoption of a mechanism that is currently being deployed at a national scale is a unique opportunity to analyze what keeps users from adopting an alternative means of online authentication after a dependable, privacy-preserving and secure infrastructure was created by the government and to reveal the social as well as technological factors that influence the lack of adoption.

For the new German ID card and its eID functionality, Poller et al. [131] postulate one main obstacle for adoption: an imbalanced cost-benefit ratio for both businesses and end-users causing a chicken-and-egg problem. Without useful and relevant services, users shy away from investing both time and money into a new technology. Yet, without a significant user base, service providers do not implement and invest in a new technology that would only replace existing mechanisms, which currently fit their needs well. Poller et al. hence believe that eID functionality is currently only viable for eGovernment applications, online applications with a traditional need for strong identity verification (e. g. banks and insurance companies) as well as ambient applications, such as age verification at cigarette vending machines. The technology acceptance model of Herath et al. also suggests potential for a lack of adoption: users may decide that their internal mechanism (i. e., the existing username-and-password practice) is suited to cope with the given situation and that the cost of adopting a new scheme is too high, which negatively affects the appraisal of a new external mechanism.

Beyond this problem, the results will show that there are further factors that influence the appraisal of an external authentication mechanism, including governmental involvement, social factors and comfort or a feeling of control. In this study, I address the following questions from a user's perspective and also present the providers' views of a novel authentication technology:

1. Which technical and social factors influence authentication behavior?
2. Which deficiencies do users see in their current authentication practice?
3. How do users perceive alternative authentication mechanisms?

[8] http://www.personalausweisportal.de/SharedDocs/Pressemitteilungen/DE/2011/Jahrestag.html

4. Why do users not adopt the available eID mechanism?
5. How does the official nature of the identity card influence their perceptions?
6. Which types of services can benefit from using eID?
7. Which measures might increase eID adoption?

First, I invited several groups of users to discuss the issues outlined above. Second, I also interviewed service providers on their perceptions of eID authentication to see which factors may motivate the use of an alternative means of authentication from the business perspective.

Overall, I found that both, users and providers, currently do not see the necessity to change their authentication methods while acknowledging problems with the current practice. Among other factors, users demanded that novel methods need to be as simple and transparent as possible, offer similar control and privacy as username and password, and are comfortable to use in all usage scenarios. In the following two sections, I will detail my findings for users and providers respectively.

3.5 The User Perspective

eID adoption has been very slow in Germany from the first day of its introduction. While many people have received a new identity card, only few services offer eID authentication. Due to the users' lack of practical experience, I chose to conduct focus groups to explore the perceptions of eID technology from the user perspective.

3.5.1 Method

Focus Groups are a variation of a group interview that "collects data through group interaction" [123]. Focus groups have been repeatedly used to "study perceptions, thoughts, and emotions" [1, 109]. Even though the use of focus groups for HCI research has been subject to discussion [138], they can extract information where other qualitative tools fail [123]. I chose this method to collect a number of factors that might influence participants' perceptions of a rather unknown topic. Since eID mechanisms are not part of common knowledge or commonly used yet, a traditional interview might have intimidated participants. Krueger and Casey [106] suggest that interviewing participants in a homogeneous group can elicit more open and honest responses, since participants realize that others also do not know so much. Additionally, discussion and interaction between participants can raise points that would otherwise not have been addressed.

I stress that the focus groups yield purely qualitative results from which I aim to extract a set of possible factors, perceptions and influences. While group dynamics may have biased the views of individual participants, I argue that I present a valid set of issues that influence the adoption of authentication technology. Furthermore, due to the nature of focus groups, I do not analyze individual views or draw quantitative

results from this analysis and will therefore not report counts for the issues raised [106].

I invited 971 students from a university mailing list to he focus groups using a screening survey that collected demographics, technical experience and Westin's Privacy Index [108]. The invitation advertised a "group discussion on daily usage behavior on the Internet" that would last for 90 to 120 minutes and offered 20 Euros of compensation. Students of computer science, information technology and electrical engineering did not receive the invitation in order to prevent anyone from disrupting the groups by being perceived as experts. This is an important consideration for focus groups, especially when discussing a topic where most people would readily defer to experts, since a single individual with extended knowledge can diminish the variety of responses and make participants reluctant to offer own explanations [106]. I received 76 complete responses to the screening survey. According to indicated time preferences, I randomly assembled 4 groups of eight people and sent out invitations for the first three. I ran three of the four planned groups before the collected views reached saturation (i. e. there were no more new aspects discussed in the third group). Participants of the fourth group as well as the remaining respondents of the survey nonetheless received a thank-you email.

Group #	**N (female)**	**Age (sd)**	**Tech. Exp. (sd)**	**Privacy**
1	6 (4)	23.5 (1.4)	3.0 (1.1)	2/4/0
2	7 (3)	24.1 (4.9)	3.6 (1.3)	2/5/0
3	5 (1)	23.0 (3.2)	3.2 (0.8)	4/1/0
Overall	18 (8)	23.6 (3.4)	3.3 (1.1)	8/10/0

Table 3.1: Overview of demographics for focus group participants. Technical Experience is rated on a scale from (1) "I often get help from others" to (5) "I often help others". The Privacy column shows the counts for the three categories of the Westin Index: Privacy Fundamentalists, Pragmatists, and Unconcerned.

Of the 24 people invited to the three groups, 18 attended. A demographic overview can be found in Table 3.1. Eight participants were female, four of which were in the first group, three in the second and one in the third. According to their Westin index, none of the participants belonged to the privacy unconcerned category, which includes individuals that do not feel that their privacy is threatened by current practice and that the benefits of disclosing data outweighs the potential dangers. Most belonged to the pragmatist category, that would "weigh the potential pros and cons of sharing information" [108]. The remaining eight participants belonged to the so-called Privacy Fundamentalists, who are most protective of their privacy. I accept the lack of unconcerned participants for the study, considering that privacy pragmatists and fundamentalists represent those groups that would express most concerns on using eID.

I moderated the focus groups together with an assistant. The moderator actively engages with the participants while the assistant monitors audio recording and takes notes. When participants arrived for my focus groups, they signed a consent form

for the audio recording, which was also announced in the invitation email. I tried to create a comfortable atmosphere, using a small conference room and unobtrusive recording equipment and also offered some refreshments and snacks. I used informal language and first names during the entire process and encouraged direct exchanges between the participants. Name tents with first names were placed on the table to increase direct interaction.

During discussion, I interfered as little as possible: I steered the conversation towards the topics of interest and encouraged participation from less active subjects. I had also prepared a questioning plan (see Appendix A.1) that gradually led participants from their current behavior and use of authentication mechanisms towards their attitudes and perceptions of eID technology in the new German ID card. I would move from one item to the next when discussion subsided or went off track. In the debriefing sessions, participants unanimously reported that they perceived the group as a non-threatening and interesting experience. Some participants indicated that they wished that they could participate in such groups on a regular basis to learn more about their own online security.

3.5.2 Results

The three sessions each lasted between 96 and 115 minutes. The audio recordings were transcribed and statements subsequently assigned to the general questions introduced above. To present the results, I report statements from all three sessions grouped by these questions. Participants are referred to as P1 to P18.

Current Authentication Behavior I began the discussion by asking for the participants' general habits on the Internet, before focusing on their authentication practice. Most participants stated that they use two to ten passwords, assigned to "service categories". The categories having the strongest passwords were often linked to attributes such as "important", "serious" or "official".

Online banking was treated with particular care by participants: they use unique and longer passwords that they generally do not write down. Participants also reported that they may actively hide security tokens used for online banking, because, in their opinion, usernames and passwords are easily obtainable by attackers. Generally, participants reported that the consequences that might arise from a compromised account affects how they choose their passwords.

Next, I asked for password management behaviors. Those who had few passwords mostly kept them in memory, while those with a larger number often used a paper-based list. Mixing passwords and usernames or using password patterns while changing individual components was also mentioned: "I have five to six passwords, two to three usernames and I mix those and then I have to remember that. Very easy. It's good for your memory [and] good for your security" (P9).

Participants also used password managers and also synced them between devices using cloud-based or personal infrastructure (USB, private server). Interestingly,

many participants did not know of password managers before the discussion. P16 indicated that he has a password manager which is secured by a fingerprint reader on his laptop, but doesn't use it because he does not care enough about security.

Registering for services was described as annoying and unnecessary at times and participants reported that they disclosed personal details only when necessary. Some participants however stated, that they did not really see any potential damage from providing their name and date of birth, for example.

Some participants argued for the need to share passwords, for example for situations where they cannot personally interact with a time-critical service. Others rejected that need entirely and stated that their passwords belong to them personally only.

Deficiencies of Current Authentication Practices When queried how secure participants feel with their current practices, participants generally said that they were confident with their schemes and that they did not see immediate problems. They offered several justifications: friends and family manage their passwords in a similar fashion and have not had problems thus far; it is too frustrating to forget a password; too many passwords get confusing or hard to manage and fewer passwords help to keep authentication fast and effortless. P6 said: "It's not that I have to think about my password, instead it simply comes without thinking [...] I type it and then I am already logged in."

Yet, a few deficiencies of current practice came up as well. Participants mentioned that password recovery fails if the registered email address is no longer accessible. P17 reported that she does not feel safe anymore, even though she increased her password strength after being hacked. In later stages of the discussion, a number of participants seemed to realize that they might not be as safe as they thought. P3 said: "you get used to it, it works well, it's easy and you stick with it. Maybe until you have a bad experience." And P10 first stated: "I think [I'm safe]. Because, someone would tell me [if I wasn't]". However, at a later point, she said "I don't feel safe anymore with my two passwords". I suspect that some participants needed external motivation to think about their online security and did not do so before participating in the focus group. But, at the same time, some participants also offered that "when I get back home, I will be too lazy [to change anything]" (P18).

Perceptions of Alternative Authentication Mechanisms I went on to ask about alternative authentication mechanisms, including using password managers and Facebook's OAuth (described by the moderator as the "Login with Facebook"-Button).

Password managers were perceived to be too complicated: "I wouldn't know how to use such a thing" (P10). This mirrors the fact that few participants had used or known about a password manager before. Participants also mentioned that saved passwords implied lost control: They stated that if someone had access to the password database, that person could get into all the services contained in the database, for example by "hacking" or using the computer when the password database is unlocked. P2 said: "I don't use it, because this way I still have, no matter whether

or not it's actually true, the feeling that I am still in control, when I really log in and that it is not an automated process." On the other hand, other participants said that they value the comfort offered by password managers higher than avoiding possible threats.

Participants also criticized being dependent on a password manager. They were afraid that they may not be able to access accounts from other locations or that they may forget the master password and therefore loose access to all accounts. Additionally, participants felt that a cloud-sync feature is unsafe because passwords are transmitted over the Internet. They also added that passwords can be compromised when the password list is on a smartphone that gets lost.

Using Facebook's OAuth had participants afraid that their information is shared with Facebook, because they felt that "they already have enough access to many things" (P11). Other sites were described to be different from Facebook and therefore there should be no interaction between those two. Meanwhile, in some cases the feature was accepted as there was no perceived threat:"I do this with Twitter, because at Twitter, there is almost no information [about me] except my email address" (P13).

Participants also said that another mechanism is unnecessary since passwords work fine. They also doubted the mechanism's security, because they don't understand what happens behind the scenes. P12 said: "Somehow I don't like those API-based logins on other sites. Some token will be passed back and forth. I don't know where it is and who does what with that. I find that suspicious." Interestingly, P12 also indicated that he manages his own server to sync his password manager. So while he knows what an API-based login is, he did not care enough to understand the workings of this carefully designed cryptographic protocol and did also not trust its creators. Finally, a loss of control was also mentioned, since Facebook might lock users out or go out of business.

Overall, participants expressed a general reluctance to adopt new services or technologies on the Internet, due to a feeling of insecurity and negative reports in the news. They showed no interest or motivation to gain an understanding of a new mechanism. P1 said that she would rather not use something "simply because then I can have a bit more security for myself". Others believed that they stick with their mechanisms because this is what they are used to and that they might be using other mechanisms, such as password managers, if they had been using them "from the beginning". Participants stated that they would wait for a mechanism to gain popularity, especially with their friends or family, before switching. They were also not ready to relinquish any comfort or mobility offered by their current practice.

Barriers to eID Adoption After discussing these more or less well known password alternatives, I introduced the eID functionality of the nPA and stated that this technology might be able to comfortably fulfill their authentication needs. Participants were told that given the necessary hardware and adoption by service providers, one would simply need to hold the ID card to a reader and enter a PIN to be securely

authenticated and that it was even possible to achieve this without disclosing any personal data using the pseudonymous identity functionality. Additionally, I stated that this would generally be more secure than using passwords, that the system is backed by the federal government and that service providers need to demonstrate a need for every piece of personal information before being granted access to this information on a cryptographic level. In order to keep the introduction short, I used simple terms and examples. Participants had a chance to ask questions in order to gain a basic understanding of the technology. Comparing different ways of describing eID functionality was not a goal of this study, as I only intended to assess the participants' perceptions of this new technology. However, if this technology were well received, participants would have already known about its features before coming to the focus groups.

After answering all questions on the eID technology, I elicited participants' attitudes towards this means of authentication. They saw the potential of this mechanism, even though most of them had not previously heard of all possible use cases. In each group, there was at least one person who had already received the new ID card. However, especially the pseudonymous identification functionality was not known to any of the participants.

When thinking about using their nPA for authentication, participants struggled to judge the mechanism because they did not know anyone using it, even though 10 to 15 million of German citizens have already received the nPA. Therefore, participants offered: "one would need to wait and find out whether or not it makes things easier and quicker" (P13). The following issues were raised during the discussions.

No added value/no motivation: Participants did not know of any relevant services offering authentication with the nPA, hence they saw no obvious advantages. Additionally, there was no motivation to adopt the new mechanism:"Honestly, I can't be bothered to look into [eID-based authentication], because I am happy with the way it is" (P10). Participants also didn't know of any services that cannot be used without the nPA.

Complexity: Participants stated that they would need a person they trust to tell them what it does and to convince them that it works. Those participants do not think that they can make that judgment by themselves. Participants also stated that they found the mechanism to be complicated. Participants said they would trust in expert reviews in computer magazines or similar reporting as well as positive experiences of family members, friends, or colleagues.

Control: Participants mentioned a fear that the system might behave in an unexpected way and that the user cannot react in a timely fashion: "[...] I'd rather have everything in my own hands" (P11), "I might forget to do something, to uncheck a box [...] I'm afraid of my own negligence" (P18). Participants also stated that they cannot be sure which information is actually transmitted. Two more technically adept participants said that they would be able to put more trust into the system, if there were independent and open-source implementations of the necessary software available.

Comfort: Participants suggested that fetching the card before being able to authenticate might be harder than relying on a password manager or one's memory. Participants were ready to make the extra effort if they saw an improvement for their security or if they do not have to remember passwords anymore. Additionally, the current need to have a dedicated card reader was also repeatedly mentioned as a barrier to adoption. Participants generally valued the comfort of using smartphones, tablets and laptops anywhere very highly and objected to the idea of reducing that comfort for purely security-related reasons.

Insufficient information: Participants who already had received the new ID card reported that the person at the public service office was not able or willing to convince them of the advantages of using eID functionality with their nPA. However, this is a crucial moment for the adoption of this technology, since users are asked if they want to deactivate the functionality when the card is picked up. Similarly, participants said that they would not be willing to adopt the eID functionality until friends or colleagues have had some positive experiences with the technology.

Mobility: most participants were not ready to carry a card reader around, but found the general idea more attractive if the card could be read with their smartphone without additional hardware. One participant also added that he disliked that the nPA is contactless and that he was afraid of being tracked or having his data sniffed without him having a chance of knowing.

Cost: Participants stated that the card readers are too expensive and offer too little added value to justify that cost at the moment.

Influence of the Official Nature of the Identity Card Participants found that a national ID card is one of the most important documents one has and it is perceived to be "a very personal document" that might not be suitable for "playing around on the Internet" (P4). They also stated the possible contradiction between being pseudonymously authenticated while using an ID card with their photo on it. Participants said, that the card is already important enough and by using it for online authentication as well, potential trouble increases when the card is lost. They also stated that they would be reluctant to use an official identity card for every-day purposes, because, in the worst case, "the government is the one that is able to storm into your house at 5 a.m. with machine guns" (P8).

A similar issue was also raised, hinting at the possibility of distributing "official spyware" in the software necessary to use eID with the national ID card, since the prevalent client-software currently comes from a government-funded project and computer surveillance by national agencies has been in the news repeatedly during the past year. A few other participants harbored similar prejudices, which were mostly based on half-knowledge. They believed that most government-funded projects are more of a failure than a success and generally do not meet deadlines or goals.

On the other hand, participants also stated that the official nature actually makes the system more trustworthy for them. One reason given as an explanation referred

to the immediate uproar in the media, when a government project has problems. Participants also felt that companies, such as Google or Facebook, can get away more easily with morally doubtful practices. On the whole, participants attributed less motivation for gathering personal information to the government than to companies.

Potential eID Use When participants where queried for which services they would more readily use eID-based authentication, they stated that they would use their ID cards on services with "an official character". This includes eGovernment services, eBanking and (health) insurance companies. Participants felt that the institutions behind these services are "more tightly bound by legal regulations" (P4) or more personal, because users know where they reside or because they have had a face-to-face encounter with someone from these institutions before. Additionally, participants would be willing to use eID with services "that already have most of [their] information anyway" (P18), such as eCommerce sites. Less "official" or less important services, such as Facebook or Skype, caused more reluctance, since these can already be used more or less anonymously if desired: "For everything that concerns my personal life and that is fun or offers entertainment, [...] I don't find [authenticating with my ID card] very useful" (P14). Participants mentioned that if they were using eID with their nPA for some services regularly, they would probably use it for all of their services.

The possibility to increase security through stronger authentication and identity assurance was acknowledged by participants. They stated that using your ID card would enable them to prove "that it is really me" (P14). However, the discussion on the utility of eID technology for different services showed that many participants, including those who had already received an nPA, had not fully understood the concepts behind eID authentication. The nPA's eID functionality was mostly reduced to how ID cards are currently used and especially the privacy-preserving pseudonymous identification functionality was quickly forgotten during the discussion. When I reminded participants of that possibility and reintroduced the concept, participants would often not see an immediate advantage in the light of other issues (see above).

Measures to make eID more attractive Towards the end of each session, I asked for suggestions to improve the adoption of eID features in the new identity card. Participants offered that "if, at some point, almost everyone used [eID-based authentication], then this might mean that it works [...] that it comes with a certain level of security" (P3). Participants said that testing the process and getting hands-on experience might help them to appreciate it. Participants also wished for proper and understandable information on this topic as well as having a competent person to talk to about the implications of using eID features. As noted above, participants that already had an eID-enabled card did not receive any guidance on the features and benefits of their new identity card, even after explicitly asking questions at a public service office. The clerks at the public service office seemed uninformed and disinterested. Participants added that banks generally explain the security measures

for online banking at length and that they would appreciate such a practice for eID functionality as well.

Generally, an increased public presence and more active marketing were mentioned as possibilities to increase public awareness. Furthermore, participants postulated that services that offer benefits through using eID might make the system worthwhile. Participants felt a need for information that was not satisfied by current practice and said they would expect television, newspaper and magazine reports about a beneficial technology. They also proposed dedicated informational events, that offer opportunities to ask questions and discuss possible uses with peers.

Additional Issues During the discussions, participants expressed that they treat the Internet as a generally insecure medium and that they therefore, for example, do not use online banking at all. Among other comments, P5 believed that password managers "surely could be hacked by someone". P9 said: "I don't believe that there will ever be perfect security on the Internet. Whether you use [an alternative mechanism] or continue using passwords [...] there are vulnerabilities everywhere." Another participant believed that there will be a way to circumvent any security system at some point in time.

We suspect that participants were not ready to invest in additional security for their Internet conduct because they don't see that this will have personal benefits in the end. Also, they might not see that security consists of several independent parts and that increasing security for one of those parts might make them safer. They do not differentiate between security risks occurring because of, for example, authentication mechanisms, lax privacy policies or missing transport security. It appears that, in several cases, this is all simply attributed to the generally unsafe Internet.

Information in Public Service Offices Because participants stated that they were not able to obtain enough information from public service offices, a colleague and I visited three of these and acted as if we were unsure of whether or not to switch to an eID-enabled ID card. During our visits, we had similar experiences as our participants: clerks were not able to answer our questions or, in one case, even refused to, saying that she was the wrong person to talk to. Yet, she was also not able to name a person to contact on this issue either. We were always referred to a brochure with a phone number inside. We tried calling that number and finally got qualified answers to the questions a layperson might have. This premium rate phone service of the Federal Ministry of the Interior can cost up to 42 Euro-cents per minute.

Summary Overall, the focus groups identified a number of problems for the adoption of new authentication mechanisms in general and eID-based authentication in particular. The factors identified during this investigation also indicate that there are barriers to adoption beyond the chicken-and-egg-problem suggested by Poller et al. [131]. The results of the focus groups allow a more detailed understanding of the

external mechanism appraisal factors presented by Herath et al. [83]: Complexity at a technical or process level and a reluctance to find out more can lead to a perceived loss of control, a lack of understanding and hence decreased motivation for adoption. Participants repeatedly stated that they do not understand technology and have no interest in it either. Hence, in order to promote privacy-preserving authentication technology to users, several other factors need to be considered before privacy benefits are appreciated. For example, participants valued comfort and mobility highly and were not ready to relinquish any in order to gain security or privacy.

3.6 The Business Perspective

The focus groups confirmed the problem of a lack of relevant services that either offer an added value or demonstrate how the nPA can make daily life easier. For trans-national companies, such as Amazon, Facebook, or Google, there is obviously little reason to adopt a technology that is currently limited to a small portion of their customer base. Yet, if many governments were to agree on a global eID standard, identity cards could be used throughout the Internet. Today, there still are several national businesses that could benefit from eID technology. Banks, insurance companies and eCommerce providers would have a means to reliably establish a customer's true identity and almost any service that has a login functionality could offer an optional eID-based authentication to appear innovative or increase customer comfort, security and privacy.

According to the list of authorized services[9], 85 public offices, companies and other businesses have been certified to access eID functionality on the nPA, of which 45 actually publicly offer eID authentication in their online processes[10]. Eight of these are not related to eGovernment applications, banks or insurance companies, which traditionally need reliable identity validation. Of these eight, only two do not request any personal information and rely on the pseudonymous functionality. The remaining 40 service providers with authorization certificates would be able to offer and use eID services but chose not to. Of those, only three sought privacy-preserving functionality, such as anonymous age verification, while the rest requested the authorization certificate to reliably establish users' true identities.

This indicates that, from the business perspective, a large number of service providers see eID features as a means to establish customer identities or to fulfill legal requirements. Using eID features of the nPA as a general means of authentication, especially in its pseudonymous form, is only adopted by two service providers after two years of being available.

[9] `http://gsb.download.bva.bund.de/VfB/npavfb.pdf` – last access: 20.9.12
[10] `http://www.ccepa.de/onlineanwendungen` – last access: 20.9.12

3.6.1 Method

In August 2012, I sent emails to 51 service providers on the list of authorized services to find out more about the service provider's reluctance. I left out infrastructure providers that mainly obtained authorization for operational or testing purposes. Additionally, I cherry-picked 26 well-known Internet services that reside in Germany and have a primarily German audience, but have not requested or received an authorization certificate yet. I asked for an interview partner that could comment on the use of eID technology in their online services. 15 providers responded, including four companies that did not appear on the authorized services list. Of the 15 respondents, two were banks, two were insurance providers, one offered free-to-play online games, one was a consulting firm that offered brochures behind a login, one a mobile phone network operator, one offered cloud-based end-user security solutions and the remaining seven were either local administrations or communal service providers for local administrations. It is important to note that eight respondents had participated in an official application testing call, run by the Ministry of the Interior prior to the nPA roll-out.

The interview partners introduced themselves as upper management, project leaders or public relation officers. Respondents from public offices were all representing the respective eGovernment bureaus. The semi-structured interviews were conducted over the phone, lasted an average of 24 minutes (sd = 11.1 min, ranging from 7 to 51 minutes), and were recorded with the interviewees' consent. Again, I was interested in extracting a set of factors that play a role in their decision making process. I extracted the central statements of each interview into an analysis sheet during a subsequent replay of the recording. Similar to the focus groups described above, I did not intend to capture quantitative data, but only to list the perceptions and evaluation results of these service providers.

3.6.2 Results

My results confirm that there is little motivation for adopting eID features through non-governmental providers. Two insurance companies stated that they provide nPA-based authentication for marketing purposes and to appear customer-friendly. The two banks I spoke to stated that the functionality currently offered by the nPA does not suffice to replace the systems currently in use to authorize bank transactions, due to the lack of a qualified electronic signature (QES). While the nPA is prepared to support QES, QES certificates have not been pre-installed on the cards by the government and could also not be purchased by customers at the time of the interviews. Furthermore, strict regulations for financial transactions and bank processes require identification that cannot be provided by the nPA. German banks have also been issuing smartcards for home/online banking since 1998. The two interviewees representing banks referred to the adoption of those smartcards and stated that, while providing considerably stronger security features, these cards were never widely adopted (cf. also Section 2.1.1). When presented with an (almost) free

alternative having lower security, private bank customers usually opt for the lower cost, according to the interviewees' statements.

Public administrations generally saw eGovernment as a necessary tool for the future, to streamline administrative processes and to offer convenient services for citizens. Since the new ID card was introduced with eGovernment as a central focus, its features are suitable for these applications. Yet my interview partners indicated that, even for eGovernment, several regulatory hurdles still need to be addressed in order to be able to provide more processes online, that currently still require citizens to visit offices in person. While two of the responding administrations were in the process of actively promoting the benefits of eID technology and the nPA, others were waiting for more adoption in the public or stronger internal demand before committing to the technology.

Many of the respondents stated that they saw problems for user adoption due to expensive or bulky card readers, lengthy and complicated user authentication procedures as well as an insufficient UI in the current version of the eID software. They also saw a need for killer applications, that demonstrate the benefits and a relevance for day-to-day use. Those who offered eID-based authentication treated this mechanism as an optional offer, that a user can but does not need to use. Obtaining an authorization to use eID features did not cause any problems. As a side note, some respondents also gave accounts of trying to increase authentication security by dictating stronger password requirements. These restrictions soon needed to be reverted since customers started to complain.

The general idea of eID technology, being able to prove actual identities in the digital realm, was welcomed by all respondents. Many stated that they expect many day-to-day processes and interactions to take place on the Internet in the future and that there is a need for effective identity management. Some respondents also acknowledged that the government can effectively roll out such an infrastructure and that users will trust in such a system eventually. Others, however, were skeptical whether or not a fear of surveillance will keep users from trusting the eID features in a personal identity card, considering computer surveillance and telecommunications data retention laws being controversially discussed.

3.6.3 Summary

Interviews with service providers showed that card features enabling providers to offer additional functionality online are either not available yet (e. g. QES) or do not meet current legal norms. Companies see little need to replace existing mechanisms beyond legal requirements to establish a customer's true identity in some cases. This is especially true since providers indicated that their customers are happy with the current practice as well. The provider interviews also confirmed the chicken-and-egg problem. Without useful and relevant services, users will not adopt a new authentication mechanism, and without user adoption, service providers will not invest in the mechanism. For the future however, service providers indicated a need for reliable digital identity management.

3.7 Discussion

With this study, I found that participants struggled to fully appreciate the benefits of the nPA's privacy-preserving features. In addition to the findings of Wästlund et al. [157], I also found that several other factors play an important role for the users' adoption and can actually overlay the perception of features beneficial for security and privacy. Overall, the results show that, in order to deploy novel authentication systems on a large-scale, effort particularly needs to be invested into service, marketing and guidance for users. This is true for both existing as well as for new systems. Surprisingly, after spending a very large amount of money on deploying an eID scheme, these factors were neglected by the German government and consequently the beneficial technical properties, such as the privacy-preserving nature of the card's authentication facilities, did not receive public attention.

Getting users' to adopt new authentication technology is also an important precursor for adoption by businesses: the interviews showed that many companies are ready to adopt new technology in order to satisfy the customers' demands. Yet, legal hurdles and insufficient technical features can also keep providers from adopting authentication technology. Similar to the users, enhancing their customers' privacy was not a central concern when evaluating a new authentication mechanism. Generally, daily use of the nPA would also benefit the government itself: Using the reliable identification provided by the card, eGovernment applications can make many processes in public administrations more efficient.

In terms of Herath et al.'s acceptance model for security technology, the results show that under the given circumstances, users do not see a problem with their current authentication method (internal mechanism appraisal). Even though the new mechanism does offer beneficial properties (external mechanism appraisal) with respect to ease of use (e. g. no more need to remember passwords), usefulness (e. g. no need to remember many passwords), and privacy, other factors tip the scales in favor of the existing mechanism. Users indicated that perceived relevance, complexity and control of the mechanism as well as cost, comfort and trust in the system play an important role when judging a novel authentication technology.

3.7.1 An Extended Model

As introduced above, a user's motivation to adopt a mechanism depends on three main factors in the technology acceptance model of Herath et al.: threat appraisal, internal coping mechanism appraisal and external coping mechanism appraisal. This means that the user needs to decide to what extent a threat applies to him or her, evaluate whether or not he or she can cope with this threat using existing – or internal – mechanisms, and whether or not an external mechanism is suited to cope with this threat. The appraisal of the external mechanism mainly depends on the two main factors outlined in TAM: Ease of Use and Usefulness. Additionally, Herath et al. posit that response costs, like concerns about the violation of privacy, play an important role in this process.

The results of this study provided some insights into the implications of this model: First, a user's awareness of threats as well as an available alternative authentication mechanism that overcomes these threats influence motivation. Without these two factors, the intention to adopt this mechanism will be low. Next, users seek information about novel mechanisms from suitable sources, which must be available for their appraisal. If these sources are not available or offer no relevant information, appraisal of the novel mechanism will be negative due to a lack of information and trust. In the focus groups, users stated that their family and friends were not using or recommending the new mechanism and that there also were no positive reports in the press. Additionally, the model implies that an authentication method that only offers more comfort or features without addressing an immediate threat will receive less adoption intention.

Based on the results, the model of Herath et al. can be extended to include more detail, as shown in Figure 3.1. I explicitly include the complexity of an authentication mechanism and the control over one's information as contributing factors into the model. Users repeatedly stated that they feel in control when using usernames and passwords from memory and that they do not understand password managers or eID mechanisms and are not interested in learning about the details.

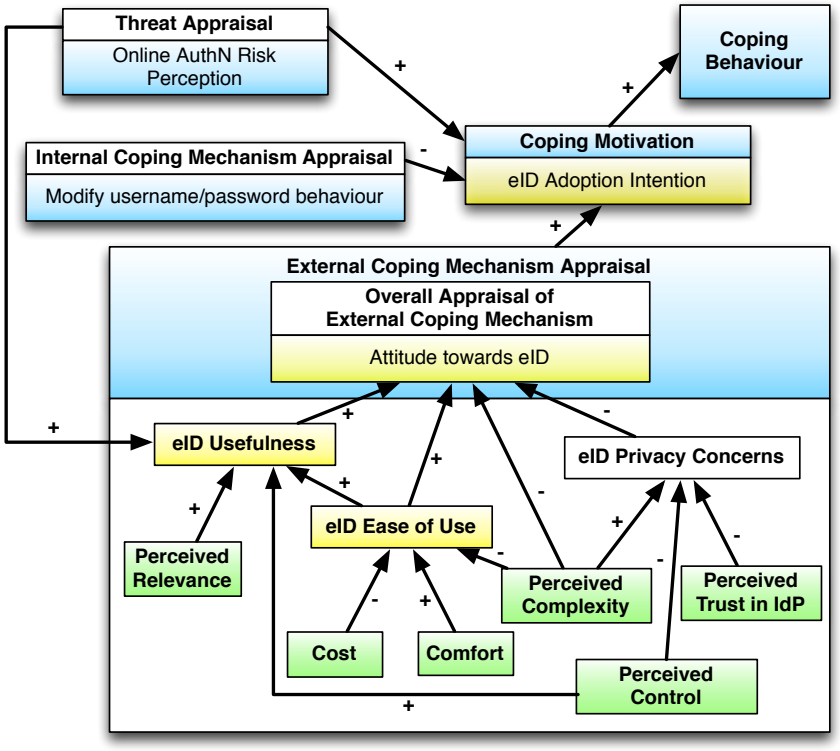


Figure 3.1: Summary of the results, based on the model of Herath et al. – Blue items are taken from TTAT, yellow items from TAM, and green items are based on my results. Arrows indicate influence relationships and their labels the direction of influence.

I also include perceived cost, comfort, relevance and trust in the provider of the mechanism to the model, since users and service providers mentioned these factors as potentially relevant. This more detailed model summarizes the results and presents a first step to gain a deeper understanding of the contributing factors for the adoption of an authentication mechanism on a large scale.

3.7.2 Recommendations

Beyond modeling the factors, this study can contribute recommendations for future efforts to provide a username-and-password alternative. First, the roll-out of the German nPA demonstrated effects of news coverage on adoption in the first few weeks after a large-scale roll-out: During the introduction of the nPA, one of the first reports in the media covered how the system is vulnerable to an attack. During the focus groups, this attack vector always came up as an argument for the vulnerability of the scheme and eventually had to be clarified by the moderator, since it was only a minor issue arising due to the use of cheap card readers. To avoid this negative influence on the users' perceptions, the entire digital ecosystem of a novel authentication mechanism needs to be considered during its design.

The study also unveiled social influences for the adoption of novel authentication mechanisms: participants stated that they would wait until there are immediate benefits, a more wide-spread trend towards the new technology and adoption by trusted third parties, such as friends, families or experts. This goes beyond a personal appraisal of an external mechanism in Herath et al.'s model and shows that users rely on several sources to inform their decisions.

Concerning the use of an eID-based online authentication solution, trust can be bipolar due to the involvement of the government. While a share of users trusts the mechanism more, others have less trust for the same reason. Future efforts should keep users' perceptions of an infrastructure or identity provider in mind.

On the technical side, novel authentication technology needs to create an effortless integration into daily workflows. While current eID systems require card readers to function, the results suggest that this additional piece of hardware is a stumbling block. Participants repeatedly mentioned that the need to buy a card reader or to carry one around would keep them from adopting eID technology for daily use. They desired comfort and mobility when authenticating online, which current smartcard and token approaches do not offer.

I thus argue that future token-based authentication systems need to leverage smartphone and NFC technology in order to satisfy these needs. Additionally, the complex and therefore opaque nature of the technology and processes underlying the nPA's eID functionality raised the users' concerns and is in conflict with their need for control.

Another important consequence of the findings is to establish user awareness: It is necessary to make users realize the problems of the authentication systems they currently use and to stress the security benefits of a new mechanism. While security

experts know the shortcomings of passwords and benefits of novel privacy-preserving technologies, the results indicate that many users feel quite safe with their current practice of using a handful of memorized passwords. Participants also showed little differentiation in terms of risks to their security on the Internet. Without knowing how certain practices can improve their security, users will show less motivation to adopt them. Adding educational material about previous systems and current problems to information campaigns and instructing staff to provide better support is thus a simple action that can be taken to improve adoption. This is surprisingly lacking in the current practice and should be included in any future deployments. Chapter 6 presents an investigation of which risks users are currently aware.

3.8 Summary

In this chapter, I examined users' perceptions and concerns on using alternative authentication methods on the Internet. As a case study, I investigated why the German nPA is receiving little adoption as a privacy-preserving authentication technology, even though the technical capabilities are excellent. The takeaway of this chapter is that "simply" ensuring that enough users get a smartcard through a national roll-out is not enough to kick-start adoption. Non-technical and social factors, such as the availability of information and reviews, positive experiences and recommendations from friends and family, as well as services with everyday relevance are necessary prerequisites for the adoption of authentication mechanisms. These findings go beyond what was theorized in the numerous models outlined in Section 2.3.

On a technical level, the results suggest that non-intrusive technology is a central factor when designing a new authentication system. I also argue that it is necessary to find a balance between technical complexity and transparency, in order to satisfy the users' need for control. A final result of the user studies shows that users need to be made more aware of immediate problems with their current practice, since unlike "functional" technology there is a lack of intrinsic motivation to adopt new authentication technology.

4 The Importance of Small Things – How Design Decisions Influence Usability

In the previous chapter, I argued for the importance of considering the entire ecosystem of a technology when trying to get users to adopt it. As I discussed in Chapter 2, usability is one of the determining factors for adoption on which developers of a security technology have direct influence. Additionally, I argued that confidentiality mechanisms find little adoption with end-users. Hence, in this chapter, I present a study on a confidentiality technology for social networks and how seemingly subtle changes in the encryption UI workflow changed users' perceived usability.

4.1 Motivation

The usability of email encryption has been the subject of research for more than a decade. As noted in Chapter 2, Whitten and Tygar [161] conducted the first Johnny study in 1999, analyzing the usability of PGP 5, followed by the more recent evaluations of S/MIME in Outlook Express in Garfinkel and Miller's Johnny 2 study [68] and the re-evaluation of the original Johnny study using PGP 9 by Sheng et. al. [146]. However, there has not been much work on message confidentiality since. In this chapter, I evaluate implementations for message confidentiality in the context of Online Social Networks (OSN) in general and Facebook in particular as well as how seemingly minor differences in the encryption workflow influence the perceived usability. As I argued in Chapter 2, usability is a central factor for technology adoption.

The messaging facilities of Facebook are a perfect example for experimentation on this subject. In contrast to the eID technology presented in the previous chapter, message encryption for Facebook can be very lightweight and rather seamlessly integrated into users' existing workflows. Also, at the end of 2013, Facebook had over 1.2 billion users[1]. In 2010, when Facebook had only 500 million users, published internal statistics showed that more than 4 billion private messages (including chat messages) were sent every day[2]. Also in 2010, a Gartner study predicted that social networking services would replace emails as the primary vehicle for interpersonal communications for 20 percent of business users in the near future[3]. To put these

[1] `https://newsroom.fb.com/key-Facts` – last access 12.3.14

[2] `http://techcrunch.com/2010/11/15/facebook-350m-people-using-messaging-more-than-4b-messages-sent-daily/` – last access 12.3.14

[3] `http://www.gartner.com/it/page.jsp?id=1467313` – last access 12.3.14

numbers into perspective, Hotmail, Gmail and Yahoo! Mail together were believed to have around 1 billion of users by the end of 2012[4].

While there are some solutions available to cryptographically protect Facebook conversations, to the best of my knowledge, there is no widespread use of them. Thus, the aim of the work presented in this chapter was to find out why this might be the case and what could be done to help OSN users to encrypt their Facebook conversations. The previous chapter focused on the ecosystem of a security mechanism as well as social factors influencing adoption decisions. Here, I focus on the user interface and immediate encryption workflow, assuming that the hurdles described in the previous chapter have already been passed. While mechanisms to protect email messaging could in principle be adapted to Facebook conversations in a straightforward manner, previous usability studies show significant problems with the existing email encryption mechanisms. One of the goals of this study was therefore to see if the changes brought about by the OSN paradigm might open up new possibilities for a usable security mechanism protecting private OSN messages.

To answer these questions, I conducted a laboratory study to evaluate the needs surrounding the protection of users' conversations on Facebook. I extracted the key usability features of several existing solutions for conversation encryption and designed mockup encryption interfaces based on these features. The results show that users state a perceived need for such a protection mechanism and that minor changes in the encryption workflow and UI influence perceived usability and adoption intention.

This chapter is organized as follows: First, Section 4.2 introduces related work relevant for this chapter, followed by a more detailed description of existing protection mechanisms for Facebook conversations in Section 4.3. Next, I describe the prototypes I built as mockups to be used in the lab study, which is detailed in Section 4.4. Section 4.5 presents the results of the study. Finally, Section 4.7 discusses limitations of my approach and Section 4.8 summarizes and discusses the findings.

Disclaimer: The contents of this chapter were previously published as part of the paper "Helping Johnny 2.0 to Encrypt His Facebook Conversations" at the Symposium on Usable Privacy and Security (SOUPS) in 2012 [56] together with co-authors Sascha Fahl, Thomas Muders and Matthew Smith. The results presented in this chapter were, however, primarily contributed by myself while Sascha Fahl provided an improved Facebook conversation protection mechanism named FBMCrypt that was detailed in the remaining parts of said paper. Thomas Muders and Matthew Smith assisted with conducting the studies and provided feedback on the manuscript.

[4] `http://www.email-marketing-reports.com/metrics/email-statistics.htm` – last access 12.3.14

4.2 Related Work

Beyond the study of Whitten and Tygar in 1999 [161] already discussed in Chapter 2, Garfinkel and Miller [68] built and evaluated a system based on key continuity management (KCM). Their prototype, CoPilot, addressed the problem of finding other users' public keys by automatically extracting senders' keys from incoming messages. Their study revealed that after using CoPilot for less than an hour, users generally understood the advantages of securing their emails. They found that while the KCM approach generally improved security, only a third of the participants chose encryption for confidential data and most sent information in an unprotected fashion. Some participants expected their email program to protect them from making mistakes and said that if encryption was important, a system administrator would have configured the email program to send only encrypted messages. This is a strong indicator that message encryption systems need to provide clear information about the security of the outgoing messages and apply security mechanisms automatically whenever possible [110].

Sheng et al. [146] conducted a follow-up pilot study to Whitten and Tygar's Johnny study with six novice users in order to understand the usability of PGP 9 and Outlook Express 6.0. Compared to the prior study of PGP 5, Sheng et al. found that PGP 9 made improvements in automatically encrypting emails, but the key verification process was still problematic and signatures in PGP 9 were actually more problematic than in PGP 5.

Outside of the messaging realm, there are several user studies that deal with Facebook privacy issues. Egelman et al. [53] ran a user study to examine how Facebook users cope with limitations of the Facebook privacy settings interface. King et al. [103] study the interaction of Facebook app users with the apps they use, what they understand of the apps' access and profile information exchange behavior and how this relates to their privacy concerns. Wang et al. [155] identified problems in the authentication dialogs for third-party apps on Facebook, proposed their own interface designs and conducted a qualitative study evaluating their designs.

The work presented in this chapter goes beyond the aspects addressed in these works by investigating which parts of an actual encryption UI have a significant influence on usability and therefore users' adoption intention.

4.3 Encrypting Facebook Conversations

As argued above, Facebook conversations are a valuable target for deploying a protection measure. To see which key design decisions have to be taken when deploying such a measure, my colleagues and I used Google, Bing and Yahoo (in September 2011) and searched for products which could be used to encrypt private messages on Facebook.

4.3.1 Existing Approaches

Encipher.it[5] and uProtect.it[6] were the only available products that could also be immediately installed. The discontinued product FireGPG was not compatible with current browsers[7], so I did not consider it a viable solution that normal users could currently install. There were, however, additional academic solutions that provided concepts for protecting different types of content on social networks. Each of these commercial and academic approaches is introduced in the following.

Encipher.it

Encipher.it provides a *bookmarklet* for Firefox, Chrome and Internet Explorer (IE) that is capable of encrypting text in any HTML text area. Thus, to encrypt a Facebook message, the user writes the message text into the Facebook message composer as usual and then has to click on the Encipher.it bookmarklet in the upper browser bar. Next, a popup in the center of the screen appears and asks the user to "Enter encryption key". Internally, Encipher.it uses AES [125] in Counter Mode [136] for encryption, i. e. the same symmetric key is used for encryption and decryption. To derive a secure symmetric key from the user's input, PBKDF2[8] is used. After a key is entered, the user must press the "Encrypt" button. The bookmarklet then replaces the clear text in the Facebook message box with an enciphered version that can be sent as normal with Facebook's "Send" button. Key management is left entirely to the user, which means the user must find a secure way of sharing the encryption key with the receiving party.

uProtect.it

Unlike the generic Encipher.it solution, uProtect.it is a third-party service specifically designed for Facebook. The user has to create a uProtect.it account and needs to install the uProtect.it browser plugin. Plugins are provided for Firefox and Google Chrome as well as a bookmarklet for other browsers. After the user has created a new uProtect.it account and installed the plugin, a green bar appears at the top of the browser window and asks the user to log into uProtect.it when the user is on Facebook. Subsequently, orange encryption buttons are placed next to text areas. Messages are encrypted and decrypted by pushing the orange button.

Also unlike Encipher.it, key management is handled automatically by the service. Unfortunately, uProtect.it does not provide any information concerning their internal security mechanisms. They do however state that they store the user content on their servers alongside the encryption keys. Thus, they are able to eavesdrop on the users' data, as stated in their Terms of Services[9].

[5] `http://encipher.it` – last access 8.4.14
[6] `http://uprotect.it` – last access 7.3.12
[7] `http://blog.getfiregpg.org/2010/06/07/firegpg-discontinued/` – last access 8.4.14
[8] `http://www.ietf.org/rfc/rfc2898.txt` – last access 8.4.14
[9] `https://uProtect.it/terms` – last access 7.3.12

Academic Solutions

Apart from the approaches above, which the average user can easily find on search engines, there are also several academic solutions which propose security for Facebook conversations. Even though these publications focus mainly on the cryptographic aspects of their solutions, each is briefly outlined in the following.

In 2008, Lucas et al. [116] proposed flyByNight, a prototype Facebook app that encrypts and decrypts messages using public key cryptography. The flyByNight server handles the key management and uses its own database to store the encrypted messages. This is a standalone app which does not protect messages sent via the standard Facebook messaging center, but rather requires the user to send all messages via the app. Lucas et al. noted that usability would be an issue for future work.

Scramble! [6] is a PKI-based Firefox plugin that can store encrypted social network content either on a third-party TinyLink server or directly at the SN provider. However, as with most PKI solutions, key management is an issue, since it relies on PGP mechanisms and must be dealt with by the user. When sending encrypted content, the user composes a message with the Facebook UI and selects the text he wants to encrypt, whereupon Scramble! requires the user to choose the contacts to encrypt the content for manually. The encrypted text or a TinyLink URL is then placed into the message composer and can be sent through the regular UI.

Another approach was taken by Guha et al. [73], who use shared dictionaries to map different "atoms" of information to a similar, valid piece of information. For example, Alice's address would be randomly replaced by Bob's, according to some mapping key. Their NOYB prototype can hide the fact that content is being protected but also necessitates key exchange using email. Additionally, reusing other users' information can have privacy implications of itself.

Baden et al. [5] present Persona, a privacy-enhanced social network platform, using public key cryptography and attribute-based encryption (ABE). They acknowledge the need to integrate their new service with the popular networks and demonstrate a prototype that provides their services as a Facebook application. They argue that existing SN apps can be gradually migrated to use the Persona platform, at least for storage. Using the Facebook API, it is however not possible to access the messaging service. The Persona user Interface and workflow for sending confidential messages is not explicitly described.

Anderson et al. [2] and Dodson et al. [47] present concepts to use rich-clients as a way to improve privacy. The SN provider is reduced to a mere content distribution server while the client handles cryptography and information semantics. This approach would require a user to migrate to another SN and change the interaction patterns, which is a different scenario from that this work addresses. In a similar fashion, a number of projects (e. g. [25, 37, 152]) proposed to distribute a Social Network across a peer-to-peer infrastructure, thereby removing the risks inherent to central service providers. However, these approaches struggle to gain broad acceptance and

are often only used by privacy-aware or expert users. One of the central ideas behind this line of work was to protect existing, widely-adopted and proprietary services transparently and therefore protect a large number of users that are already subject to privacy problems.

4.3.2 Extraction of Features with Usability Impact

Unlike in the related email-based studies, where relatively mature and stable implementations of PGP and S/MIME were available and could be studied directly, the solutions for Facebook are partly general purpose encryption products which can also be used with Facebook or early academic prototypes and niche products with usability issues which stem more from implementation limitations than design issues. For this reason, I decided to extract the design decisions and build mockups to study the basic building blocks and their usability issues. A further reason for choosing this abstract approach over a direct product evaluation was that the two available solutions, Encipher.it and uProtect.it, differ in several key areas, which would have made it very difficult to judge which features made the one more usable than the other. Thus, I extracted core features of the above solutions to study the usability of encryption for Facebook conversations.

	Encipher.it	**uProtect.it**
Key Management	manual	automatic
Encryption	manual	manual

Table 4.1: A comparison of key management and encryption/decryption concepts applied by Encipher.it and uProtect.it.

Three features are particularly well suited to distinguish the above approaches: encryption UI, key management and integration. For the encryption UI, some solutions require the user to trigger the encryption process manually by activating a bookmarklet or pressing a button, others trigger encryption automatically. The different key management options require the user to get involved in the key management process by manually sharing or selecting keys, while other solutions automate this issue. A further feature is integration. Some solutions require the user to send private messages via a completely separate UI instead of Facebook's standard UI, while others integrate their solution into Facebook. In order to keep the study design as simple as possible, I chose to focus on integrated solutions, because forcing the user to leave the Facebook UI is already an undue burden on usability. Table 4.1 gives an overview of the values for the two remaining variables in the two real-world solutions. Based on this, I built four mockups, described in the following section, which were then used for the laboratory study.

4.3.3 Mockups

To evaluate the different interface and workflow concepts for sending encrypted Facebook messages as discussed above, I built mockups using Greasemonkey[10]. The mockups allowed me to test the independent variables shown in Table 4.1 in the context of sending encrypted private Facebook messages.

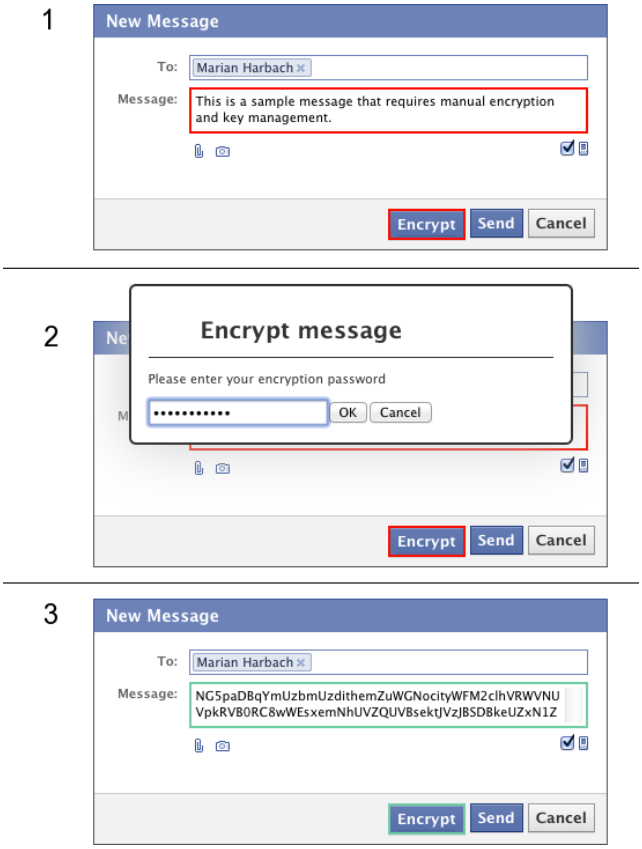


Figure 4.1: The three steps in mockups 2 & 4.

Figure 4.1 shows mockups 2 and 4 corresponding to manual encryption combined with both manual and automatic key management. In the case of manual encryption with manual key management, the user enters the message text as usual (Step 1). The user must then click the new "Encrypt" button. A popup asks the user for an encryption password with which the message is encrypted (Step 2) and the resulting ciphertext is placed in the message box. The user can then send the message using the original "Send" button (Step 3). The unencrypted text and the "Encrypt" button are marked by a red border to draw attention to the changes. Successful encryption changes the colour to green. The encryption password must be shared with the recipient manually. This corresponds to the Encipher.it workflow. All steps are repeated for every message sent.

[10] `http://www.greasespot.net/` – last access 8.4.14

In the case of manual encryption with automatic key management, the key management model from uProtect.it is used to replace the manual key management of Encipher.it. This means that Step 2 only needs to be executed once per Facebook session, since the password can be cached locally and the entered password does not need to be shared manually with the recipients.

Figure 4.2 shows mockups 3 and 5 corresponding to automatic encryption combined with both manual and automatic key management. With these mockups, the user does not need to manually trigger encryption. Rather, when the Facebook "Send" button is pressed (Step 1), encryption is triggered automatically. In the case of manual key management, the user needs to choose an encryption password for each message to be sent and share it with the message recipient manually (Step 2). In the case of automatic key management, the user only needs to enter the password once per Facebook session, as in the uProtect.it workflow. In order to offer a similar amount of visual feedback as in mockups 2 and 4, the message is not sent instantly after completion of Step 2. Instead the ciphertext is shown in the message composer's text area with a spinner animation for two seconds to visually indicate successful encryption after which the message is sent (Step 3).

Figure 4.2: The three steps in mockups 3 & 5.

I also added red and green visual security indicators to the text area and the "Send" button of mockups 2-5 as a visual aid, as suggested by Egelman et. al. [54] and Maurer et. al. [121].

In addition to mockups 2 to 5 described above, I built mockup 1 without any modifications of the Facebook message composer to serve as control condition.

4.4 Laboratory Study

Based on the mockups introduced above, I conducted a laboratory study to evaluate the usability impact of the basic building blocks of a message protection mechanism. I tested the influence of manual vs. automatic encryption and manual vs. automatic key-management on usability, adoption intention and perceived security. I also wished to find out which role password or key recovery plays for the adoption of an encryption mechanism. Depending on the cryptographic principle used, the loss of the encryption key can result in complete loss of the encrypted data in case the key cannot be recovered and as a result decreases both the acceptance and the utility of a solution.

4.4.1 Participants

To recruit users for the lab study, I conducted a screening survey. The survey was also used to gain an overview of the level of interest in protecting these conversations. I invited 16,915 students at the Leibniz University Hannover via email to participate. The study was introduced as a study on Facebook privacy and advertised 10 Euros of compensation for a 1 hour study. The invitation was not designed to hide the fact that the study was related to Facebook message privacy, since I was interested in participants who would potentially want to use an encryption mechanism to protect their Facebook conversations. Educating or motivating participants who are not worried about their conversation's privacy was outside the scope of this work.

I randomly selected candidates from the 564 respondents of the screening survey, who met the following criteria: they needed to be concerned that Facebook could access their private messages and needed to use Facebook at least on a weekly basis. I excluded infrequent Facebook users to minimize the risk of technical difficulties when using Facebook. Finally, I excluded computer science students to avoid bias based on technical skills and possible familiarity with encryption mechanisms.

This left 291 possible candidates, from whom 100 were randomly selected for the study. 96 of these attended their appointed time slot. All participants were students of Leibniz University Hannover. Table 4.2 outlines the demographics of the participants.

N	96	
Gender	44	male
	52	female
Age	22	years (sd=2)
FB Membership since	23	1 yr or less
	37	2 yrs
	34	longer
	2	don't know
FB Password lost (prev. 12 months)	79	not once
	9	once
	8	twice or more often
Facebook use	10	several times per week
	27	less than 1 hour per day
	41	1 - 2 hours per day
	17	2 or more hours per day
Facebook Friends	207	friends on avg ($sd = 130$)
Facebook Messages	24.4	msgs per week ($sd = 46.7$)
Used encryption before?	33	yes
	63	no or don't know

Table 4.2: Lab Study Demographics.

4.4.2 Ethical Considerations

The study was conducted in Germany and thus was not required to pass an IRB review. Nevertheless, the study complied with the strict German privacy regulations. I did not use the participants' real Facebook accounts and all data was collected anonymously. After the study, participants were debriefed and any questions regarding the study were answered.

4.4.3 Procedure

The study took place in a computer lab in the university, where multiple PCs with Firefox 9, Greasemonkey, the mockups and a webmail interface for each participant was set up. I created artificial Facebook accounts and email addresses, so that the participants did not have to use their real accounts and data. The mockups simulated sending private messages, rather than actually sending the messages which might have accidentally triggered the anti-spam protection of Facebook, resulting in blocked accounts. However, I did ensure that the mockups appropriately simulated the behavior of Facebook's standard message composer, so that participants would not notice that messages were not actually sent.

The participants were informed that they would be testing five different technologies to encrypt Facebook conversations. Table 4.3 gives an overview of which condition and mockup is dealt with in which task.

Task	Interface	Encryption	Key Management
T01	Control	None	None
T02	Mockup 2	Manual	Manual
T03	Mockup 3	Automatic	Manual
T04	Mockup 4	Manual	Automatic
T05	Mockup 5	Automatic	Automatic

Table 4.3: Properties of the tasks in the lab study.

To avoid bias during the study, participants were instructed that the technologies were not built by myself and that the study was testing the technologies, not the participants. Each participant was supervised by a study monitor, who measured the time needed to complete each task and noted errors. The monitor was allowed to assist with the browser tabs and the webmail program, but no help or information was given concerning the mockups or the tasks themselves. The next section outlines the basic structure of each task.

Tasks

To keep the design simple, all tasks were focused on encrypting and sending private Facebook messages to three different friends (Jan, Vanessa and Heike). The decryption process is often analogous to encryption and was therefore not tested explicitly.

Handouts were given to the participants which explained the procedure of sending an encrypted message with the given technology. The messages to be sent were as follows:

To Jan: Hello Jan. Please transfer the money to my bank account, account number 123456 and sort code 100200.

To Vanessa: Jan has transferred the money to my bank account.

To Heike: Hi Heike. Have you transferred the money yet?

Since all participants had a self-reported interest in protecting their Facebook conversations from unauthorized access, I chose sample messages which contained financial information with the aim of inducing a similar wish for privacy in all participants.

T01 was the control task. Participants were asked to send the messages using the regular Facebook message composer. The task was used to get a baseline for error rates and speed of the individual participants. Like in the other tasks, participants were told that their messages were encrypted. In contrast to T02 to T05, message encryption was not featured explicitly, but included in the regular sending process without visual indicator or actions. The control task therefore additionally lends itself to examine whether or not the participants would accept and trust a mechanism that provides “invisible” security.

During the manual key management tasks (T02 and T03), the participants needed to use a webmailer interface to send an arbitrarily chosen key to the corresponding recipients out-of-band. Using a webmailer is of course not the optimal out-of-band solution in terms of security. However, since the study's focus was on the Facebook UI and not the out-of-band communication capabilities of the participants, email was used as a mechanism which would cause little technical trouble during the study. In a real world setting additional problems could arise here.

In the automatic key management tasks (T04 and T05), only the first message required participants to enter their password. Participants were told that the password was cached for the rest of the session. This corresponds to the user logging in to a service such as uProtect.it. All participants were given the same password and were informed that *"The decryption password is automatically and securely transferred to the recipient, so that he can decrypt the message"*.

The only difference between the manual and automatic encryption tasks is that the "Encrypt" button needs to be pushed before sending the message.

4.4.4 Study Design

Since the study encompassed reading and comprehension, my co-authors and I chose a within-subjects design to compensate for individual differences [112]. To minimize bias due to learning effects, tasks were assigned based on latin squares, so that each task was equally distributed over each position in the within-subjects design. To avoid skewing results due to effects of familiarity, only participants who use Facebook frequently and are hence familiar with the Facebook UI were invited, as stated above.

In post-task questionnaires for each of the five tasks, I collected the system usability score (SUS, ten items addressing multiple facets of general system usability [24]) as well as additional items concerning participants' willingness to use the corresponding mechanism in the future for private and general messaging ("adoption intention"). A final item assessed how well participants felt their messages were protected. The contents of the post-task questionnaire can be found in Appendix B.2.

After completing the five tasks, participants were given a final questionnaire (cf. Appendix B.3). Apart from gathering demographic information, this questionnaire also presented a hypothetical question, asking whether or not the participants would use an encryption method which would render all previous encrypted messages unreadable if they forgot their password. I also asked supporting questions to ascertain the reasoning behind this decision.

4.5 Results

The results of the lab study are presented below, grouped by the research questions that drove the experimental design.

4.5.1 Desire for Confidentiality in OSN Conversations

In the screening survey for the lab study, participants were asked whether or not they thought that Facebook could read their private messages as well as whether or not this would be a cause of concern for them. I received 514 complete responses. Of these, 413 (80.4 %) were aware that Facebook was able to access their private messages. When asked whether this concerned them, 263 (63.7 %) answered "yes", 78 (18.9 %) answered "no" and 72 (17.4 %) stated they didn't care. The remaining 101 (19.7 %) participants stated they were not aware that Facebook could read their private messages. When these participants were asked whether it would concern them if Facebook was able to read their private messages, 79 (78.2 %) answered "yes", 12 (11.9 %) answered "no" and 10 (9.9 %) stated that they did not care. In total, 342 (66.5 %) of the 514 participants stated that they were or would be concerned by Facebook being able to read their private messages. Thus, there appears to be a certain awareness of problems and a potential for a desire for protection.

4.5.2 Usability

In the lab study, the central concern was the impact of changes in the encryption workflow on usability, adoption intention, and security feeling. In terms of usability, tasks T04 and T05 as well as the control task T01 received the highest mean system usability score (SUS). Table 4.4 summarizes the usability scores.

Task	SUS	sd_{SUS}
T01	88.20	15.32
T02	64.27	18.56
T03	65.86	18.43
T04	86.51	11.43
T05	89.79	14.20

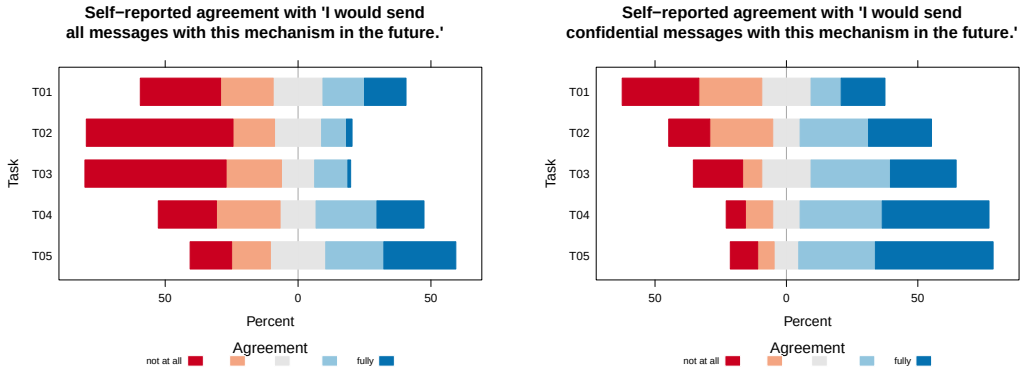
Table 4.4: Mean usability (SUS) across tasks.

The data was not normally distributed as the SUS score interval is limited at 100. The statistical analysis was hence conducted using non-parametric tests. A Friedman's ANOVA yielded a highly significant omnibus effect ($\chi^2(4) = 197.19$, $p < .001$) and Bonferroni-corrected post-hoc comparisons yielded significant differences between tasks T01, T04, T05 and T02, T03. This shows that the manual key management tasks were perceived to be less usable than the other three conditions. Also, adding encryption either manually or automatically did not matter for usability and these two conditions also performed as well as the control.

4.5.3 Adoption Intention

With respect to adoption intention, tasks T04 and T05 also received the highest ratings for both sensitive and general messaging (on a 5-point numeric scale). Table

4.5 and Figure 4.3 give an overview of participants' responses. Figure 4.3a clearly shows the same trend found in the usability results: conditions with manual key management had lower adoption intention ratings compared to the other three conditions (Friedman's ANOVA, $\chi^2(4) = 99.9$, $p < .001$, Bonferroni-corrected pairwise comparisons, $p < .05$).



(a) For all messages **(b)** For sensitive messages

Figure 4.3: Adoption Intention rating across the five tasks.

For sensitive messages (cf. Figure 4.3b), an omnibus Friedman's ANOVA also showed a significant difference between conditions ($\chi^2(4) = 50.5$, $p < .01$) but Bonferroni-corrected pairwise comparisons only yielded significant differences between the control and T04 as well as T05. This shows that participants would not regularly use the plain Facebook messaging dialog for their sensitive messages, even though I told participants that the message were encrypted.

Task	a_{priv}	sd_{priv}	a_{all}	sd_{all}	sf	sd_{sf}
T01	2.62	1.438	2.67	1.449	1.57	0.778
T02	3.19	1.439	1.87	1.136	3.49	1.133
T03	3.35	1.421	1.87	1.117	3.42	1.158
T04	3.87	1.259	2.91	1.437	3.20	1.148
T05	3.92	1.319	3.30	1.415	3.23	1.174

Table 4.5: Mean ratings of adoption intention for private (a_{priv}) and all messages (a_{all}), as well as security feeling (sf) on 5-point numeric scales across tasks.

Adding up the ratings across tasks with similar encryption modes (manual in T02 and T04; automatic in T03 and T05), a Wilcoxon signed-ranks test yields significant differences ($V = 581$, $p = .007$) between manual and automatic encryption for all message adoption intention. This effect was however not quite significant for private messages ($V = 727$, $p = .1$). It appears that when considering the use of a security measure for all messages one sends, small differences, i. e. having to click an additional button or not, influence participants' perception.

Figure 4.3 also shows a general difference between adoption intention for all messages and more sensitive messages. Participants are more reluctant to apply a protection

measure in general compared to only for specific, sensitive messages. This effect is also highly significant (Wilcoxon signed-ranks test, $V = 121$, $p < .001$).

4.5.4 Correlation Between Usability and Adoption Intention

To test the correlation between the perceived usability and the stated adoption intention of a message protection mechanism, I used Kendall's tau and found significant values in all five tasks (see Table 4.6). The correlations suggest that higher usability correlates significantly and to considerable extent with higher adoption intention.

The differences in correlations between sensitive and all message was not significant (t-Test, Chen and Popovich [27]), except for the control task T01.

Task	$\tau_{sensitive}$	p	τ_{all}	p	p_diff
T01	.199	.012	.197	.013	.003
T02	.440	$< .001$	.287	$< .001$	.76
T03	.367	$< .001$	.194	.015	.55
T04	.421	$< .001$	.325	$< .001$	.54
T05	.436	$< .001$	.413	$< .001$	.052

Table 4.6: Spearman's correlation between usability and acceptance for sensitive and all messages across tasks. p_diff denotes the p-value of the t-test between the correlation coefficients of each task [27].

4.5.5 Perceived Protection

In order to investigate the perceived level of protection across mechanisms, I ran a Friedman test on the participants' answers concerning their perceived protection in tasks T02 through T05. I found a highly significant difference in the mean ranks ($\chi^2(3) = 15.947$, $p < .001$). The top-2-box scores show that in tasks T02 and T03 54.2 % of the participants felt well protected and in T04 and T05 only 41.7 % and 40.6 % felt the same way. I therefore suspect that the complexity of a mechanism – in this case creating individual encryption keys for each recipient and distributing them manually – heightens a user's subjective sense of security. Figure 4.4 summarizes the ratings.

It is noteworthy that only 2 % of the participants felt well protected in the control task. Even though they had been told that the mechanism presented in T01 would protect their message, they apparently placed little faith in this statement. While this could be due to their familiarity with Facebook, I also suspect that an entirely invisible and effortless protection mechanism does not generate a feeling of security and is not trusted by users. This is an interesting question, since "invisible" security is often claimed to be a desirable feature. However, the results suggest that trust in the mechanism could be a problematic issue. This study was however not set up to analyze this observation further.

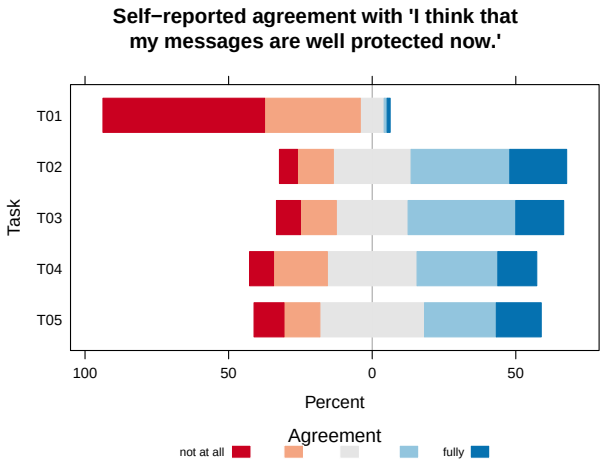


Figure 4.4: Rating of security feeling across the five tasks.

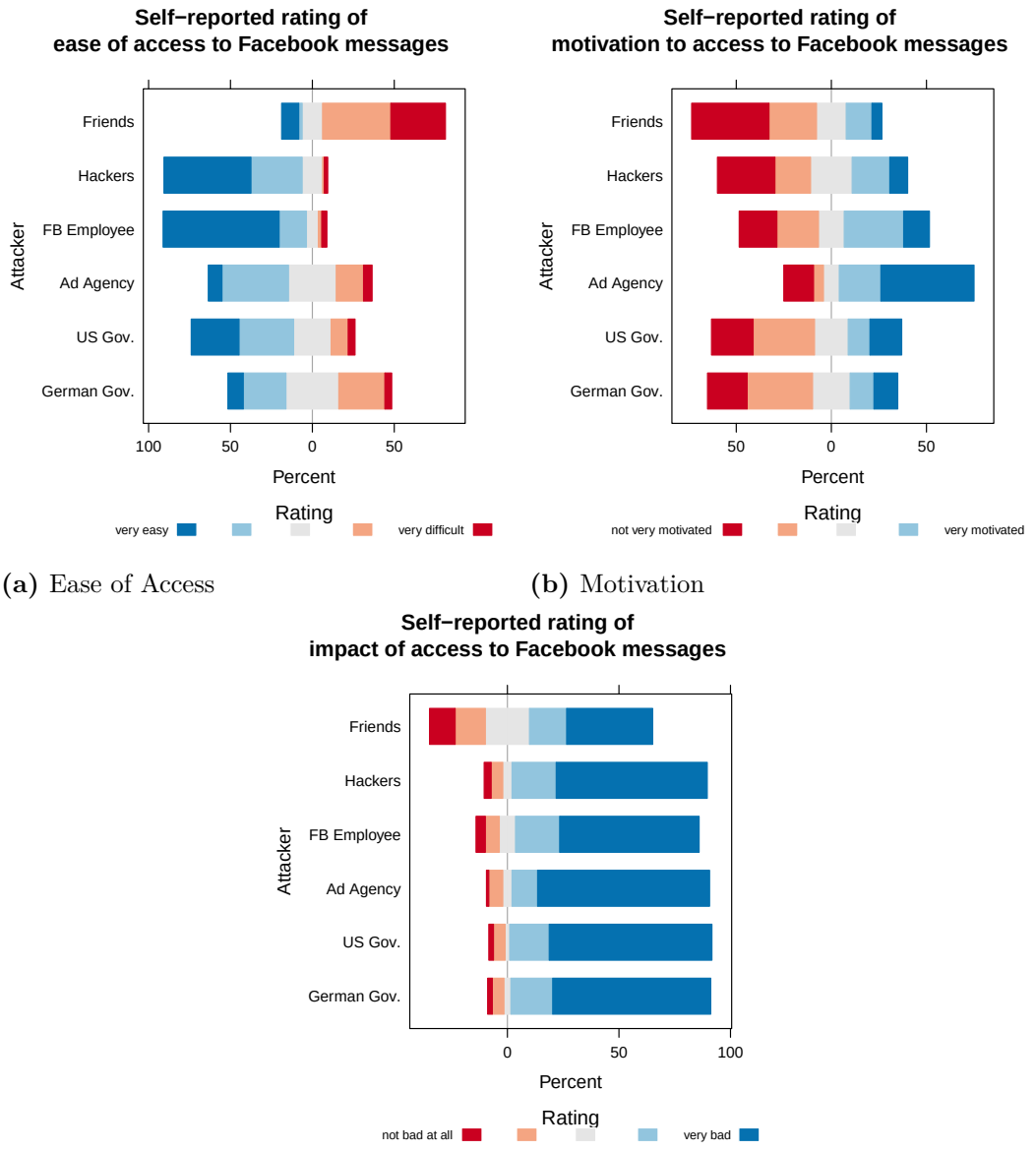
4.5.6 Password Recovery

To test how fear of losing data influences the need for password recovery, I divided participants into those who stated that they were worried or very worried about losing all their old messages or forgetting their password (group A, $n = 49$) and those who were not (group B, $n = 47$), using top-2-box scores of a 5-point numeric scale. I found a significant difference between group A and group B using a Chi-Square Test concerning whether or not they would use a mechanism without password recovery ($\chi^2_1 = 18.383$, $p < .001$) and whether or not they would prefer a mechanism with password recovery ($\chi^2_1 = 10.341$, $p < .001$). In group A, 72.3 % would not use a mechanism without recovery and 78.7% would prefer a mechanism with password recovery, while in Group B these figures were 28.6 % and 46.9 % respectively.

4.5.7 Perceived Privacy Threats

To analyze who was perceived to be the biggest privacy threat, participants rated how easy it would be for different entities to access their Facebook conversations on a 5-point numeric scale. Figure 4.5a provides an overview. Facebook employees and hackers were perceived as having the easiest access to that information: 87.5 % and 84.4 % said that they thought it would be easy or very easy for these actors to access their private messages, followed by the government of the USA (62.5 %), advertising agencies (49.0 %) and the German government (35.4 %). Only 12.5 % believed that it was easy or very easy for their friends to access these messages.

Additionally, I wanted to know how motivated the participants believed these entities would be to access their messages (cf. Figure 4.5b). Advertising agencies were believed to be the most eager (70.8 %). Facebook (44.8 %), Hackers (29.2 %), the US government (28.1 %) and the German government (25.0 %) are believed to have



(a) Ease of Access

(b) Motivation

(c) Perceived Impact

Figure 4.5: Participants' ratings concerning access to their Facebook conversations across different potential attackers.

less motivation to access private messages. Friends were believed to be the least motivated (18.2 %).

Finally, I asked how bad the participants would feel if these entities accessed their private messages (cf. Figure 4.5c). 55.2 % would find it bad or very bad if friends could access private messages not intended for them. For all the remaining actors, the participants almost unanimously agreed that access to their private messages would be bad or very bad (82.3 % to 90.6 %).

4.6 Discussion

The results presented in the previous section clearly show that a single design decision for a security measure, namely letting users handle the encryption key by themselves or providing automatic key management, had a severe impact on perceived usability and hence adoption intention. While not as large as the impact of the key management mode, there also was an almost-significant effect between manual and automatic encryption. When participants had to push an extra button, the usability and adoption intention was reduced. I also found significant correlations between usability and adoption intention, supporting that there is an immediate and strong connection between the two concepts.

The results on the perceived protection ratings show an interesting pattern: the less effort a security mechanism required and the less visible the protection mechanism becomes, the less participants felt well protected. This is the opposite trend found between usability and adoption intention. Another feature was not present in the task and mockups, but was also stated to be desired by participants. Password recovery – which is often not possible in strong cryptographic schemes such as asymmetric encryption – was particularly important for participants who were afraid to lose their password. Again, this basic design decision influences adoption intention.

Finally, the analysis of potential attackers and their motivation showed that in this case, participants mostly wanted to be protected from more or less unknown attackers which they however in several instances believed to be highly motivated to access their data. Assessing these perceived threats can help to tailor security solutions to the participants' needs. For example, if a Facebook conversation encryption mechanism advertised to fend off hackers and ad agencies, adoption intention could increase as participants will more readily associate a measure's protection potential with their own goals and beliefs.

4.7 Limitations

The work presented in this chapter has the following limitations.

Precision: due to the within-subjects design of the lab study, carry-over and fatigue effects could have affected the study results. While a between-subject analysis based on the latin square setup did not show any worrying trends, a larger dedicated between-subjects study would be needed to rule out these effects.

Generalizability: Participants were all university students, selected for their frequent use of Facebook and their desire for Facebook message privacy. I argue that the two selection criteria are valid, since this is the target group of Facebook encryption mechanisms. Additionally, university students are open to new technology and use it frequently, hence presenting a best-case scenario for adoption intention. However, future studies of participants outside the university's demographic is of course desirable.

Realism: Participants were restricted to using the computer provided for them during the study and using dummy Facebook and email accounts. Furthermore, only the first-time user experience was studied; I did not examine daily usage behavior. Long-term studies using real Facebook accounts would address this.

4.8 Summary

In this chapter, I have presented a user study on the impact of small changes within the encryption workflow for Facebook messages (automated vs. manual key-management and automated vs. manual encryption) on adoption intention and usability. The screening survey of the experiment with 514 participants showed that, within the sampled student population, there is a desire to protect Facebook conversations. In the experiment with 96 participants, I found highly significant preferences for automatic key-management and automatic encryption, indicating that making users handle their cryptographic keys by themselves and making them push a button to encrypt a message reduces usability and hence adoption intention. Furthermore, participants who were worried about forgetting their password or losing access to their previous conversations stated that they would not use a mechanism without password recovery. Even though the automatic mechanisms had higher adoption intention ratings, I also found that the two more complicated encryption mechanisms generally made the participants feel better protected. In contrast, invisible encryption let participants doubt the security of the product.

Overall, the results presented in this chapter show that it is important to design security UIs and workflows to suit the users' need for protection as well as their habits. Manual exchange of cryptographic keys lowered adoption intention significantly. Also, a lack of password recovery facilities and trust in the protection abilities is a reason to not adopt a measure.

5 Security Measures in the Wild – A Smartphone Case Study

Beyond the general decision on whether or not to adopt a security measure at all, a user's need for protection may be fine-grained and dependent on context. To investigate both, the influence of general risk perception as well as risk perception in-situ, this chapter presents results of an investigation about how users perceive and interact with a very common protection mechanism, the lock screen of their smartphone. The study uncovers which factors motivate users to (not) apply such measures in everyday life, which additional measures they take, and in what kind of situations additional measures may be desired.

5.1 Motivation

The two previous chapters looked at factors that drive a general decision for or against a certain measures. In this chapter, I present an investigation of users' reasoning about an existing measure that they are already using. While, in the previous chapters, I described how to design new security measures so that they will find adoption, this chapter is about understanding how the use of a security measure in everyday situations influences perception. Smartphones were a natural choice for such an investigation. Current mobile devices are touch-based, rich in functionality and provide high memory capacity. While early devices needed key locking mechanisms solely to prevent accidental use, current smartphones require protection mechanisms because of the potentially vast amount of private data contained on the phone. As a consequence, users may feel a greater need for protection and, therefore, authentication on mobile devices has become indispensable for many users. Besides traditional alphanumeric passwords and PINs, current smartphones also provide graphical as well as biometric authentication mechanisms. Given that users can make an ad-hoc decision to choose another lock screen and that they carry their phone in a multitude of different contexts, smartphone locking behavior is a promising subject to study in order to gain insights into the reasoning of users about the necessity of protection measures in daily life.

To address the – at least theoretical – need for better protection of private information on smartphones, research concerning mobile authentication is very active. In this area, one of the most cited dangers for smartphone unlocking mechanisms are shoulder surfing attacks (e.g. [12, 117, 142]). That is, direct observations with and without technical equipment (e.g. camera) aiming to capture a user's password.

Based on this assumption, most proposed unlock mechanisms pay particular attention to being resistant against shoulder surfing and consequently accept reduced usability (e.g. [11, 42, 101]). Based on what we have learned so far, this may severely impact adoption, if users do not agree that this is a relevant risk.

Interestingly, even though shoulder surfing is often assumed to be a relevant real-world problem, there is almost no data on the occurrence of shoulder surfing attacks in the wild or on users' perceptions of the threat. Furthermore, since lock screen mechanisms are often tested in lab environments, little is known about the users' perceptions and their behavior in real-world situations. Amongst others, important research questions answered in this chapter are: How often and in which situations do people use secure lock screens? How often and in which context do people access sensitive data using their phone? How often is this data perceived to be in danger? And to what extent is shoulder surfing perceived to be an issue in everyday mobile device authentication, justifying more secure but also more complex protection measures?

To shed light on these questions, colleagues at the Ludwigs-Maximilian University in Munich and I conducted an online survey (n=260) and a field study (n=52), analyzing users' risk perception and behaviors when interacting with smartphone unlock mechanisms. We gathered in-depth insights into the assessment of shoulder-surfing risks and shed light on users' perceptions and daily needs when protecting their smartphone. Our approach allowed us to provide a quantitative analysis of real-life unlocking behavior and analyze influences of perceived risk and environment.

Some of the key findings discussed in this chapter are that users spend up to 9.0 % of the time they use their smartphone on dealing with unlock screens, that a secure lock screen is considered unnecessary in 24.1 % of the situations we sampled, and that shoulder surfing is only perceived to be a relevant risk in 11 of 3140 sampled situations. We also found a very diverse set of justifications for (not) having a secure lock screen, a plethora of physical measures users take to protect their phone, and that losing the smartphone-hardware is the most relevant threat to users.

The understanding gained from these studies needs to play an important role in the design of future unlocking mechanisms, since the usability/security trade-offs of current mechanisms do obviously not fully match users' concerns and may hence limit adoption.

Disclaimer: As this work was conducted together with Emanuel von Zezschwitz, Alexander De Luca, and Andreas Fichtner from the Ludwigs-Maximilian University in Munich as well as Matthew Smith at the University of Bonn, this chapter will use the academic "we" in the remainder of the text to mirror this fact. The contents of this chapter were also previously published and presented at the Symposium On Usable Privacy and Security (SOUPS) in 2014 under the title "It's a Hard Lock Life: A Field Study of Smartphone (Un)Locking Behavior and Risk Perception" [79]. The idea and initial concept for this work came from myself, while my colleagues from Munich provided additional input for the study design, and jointly conducted

the study together with me. I analyzed the data before we jointly discussed the implications and compiled the paper for publication.

5.2 Related Work

Chapter 2 already introduced the existing theories on the adoption of security measures and the compliance budget as well as previous work on the perception of security. There are two additional areas of related work relevant to this chapter. We will first outline the very active field of smartphone lock screen research to motivate the need for ground truth on the threats users face in their daily lives. Then, we present existing data on the use of security measures on smartphones.

5.2.1 Unlock Screens

Authentication on mobile devices can be divided into implicit (e.g. [94]) and explicit approaches (e.g. [154]). In addition, there are mixed approaches (e.g. [40]) which add implicit security layers to an explicit authentication challenge. Implicit authentication mechanisms analyze specific time spans of behavioral cues like sensor data and usage patterns to establish a continuous authentication and hence reduce authentication workload. Examples include analyzing gait patterns [149], typing behavior [32], file system access [166], or a combination of factors [135]. Due to noticeable delays, many of them are not suited for direct lock screen mechanisms. Explicit authentication methods can be divided into biometric, token-based and knowledge-based methods [127]. The latter face the threat of shoulder surfing attacks [117].

As a consequence, the goal of finding shoulder surfing resistant solutions for knowledge-based unlock screens has become a very active research area (e.g. [12, 117, 142, 150]). Proposed concepts achieve shoulder surfing resistance either by establishing secret channels [11], by utilizing indirect input [41, 42, 101, 102], by obfuscating the input [167], or by adding additional biometric layers [40, 145].

Developing usable authentication mechanisms, which are secure against attacks such as shoulder surfing is believed to be very important. Nevertheless, to date there is no evidence that the often postulated threat of shoulder surfing attacks holds true in the users' daily lives. Therefore, as shown in the previous chapters, shoulder-surfing resistance may not be a valid motivator for users to adopt such measures. All of the above works were also evaluated in laboratory settings and established concepts like PIN or patterns solely serve as a baseline. User perception and field performance (even of PIN and patterns) therefore remain largely unexplored. The only published work in this area focused on a quantitative performance analysis of PIN and patterns, but did not analyze real lock screen interactions [154].

Karlson et al. [98] already argued for better support of phone sharing through non-binary locking mechanisms. Prototypes of context-aware or selective authentication mechanisms for smartphones have also been proposed by Hayashi and colleagues

[80, 82]. They report being able to reduce the number of authentications by up to 68 %. To date, however, there is only limited data on how these mechanisms relate to users' needs during their everyday smartphone use and which factors drive users' decisions for or against an authentication mechanism. We provide further evidence for the advantages these approaches can have, not only with respect to user workload but also to reduce the attack surface for shoulder surfers.

5.2.2 Security Perception and Smartphone Use

There have been several non-academic studies that report how frequently users interact with their smartphones. For example, a study by lock screen advertising provider Locket finds that the users of their app unlock their phones 110 times a day on average[1]. In a recent market research study by Nielsen[2], researchers found that smartphone users in the UK spent almost 42 hours interacting with their smartphones in December 2013. This figure was somewhat smaller in the U.S. (34.3 hours) and Italy (37.2 hours). Phone and operating system manufacturers likely also have usage data, but these data sets are not publicly available.

5.3 Online Survey

To begin to understand how users think about smartphone locking, we conducted an online survey. The aim of the survey was to get an overview of users' concerns and motivations for locking or not locking their devices. Research questions included: Why do users (not) lock their phone? Which factors play a role in their decision making about this security measure? Which kinds of attack scenarios do users consider? Are users more afraid to lose their phone in general or that someone will actually access their data? Are there any additional measures that users frequently take to protect their phones and how do these relate to having a lock screen or not?

5.3.1 Method

We used Amazon's Mechanical Turk (MTurk) service to distribute the survey. While MTurk does not allow us to draw representative samples of any population, the people that participate in this service have been shown to generate meaningful results in the area of usable security [99] if appropriate precautions are taken [49]. We advertised a survey about smartphone use in daily life and offered $0.70 of compensation per successfully completed task. We asked participants to only take the survey if they have been using a smartphone regularly for at least three months. They had to prove their ownership of a smartphone at the end of the survey by scanning a

[1] http://www.npr.org/blogs/alltechconsidered/2013/10/09/230867952/new-numbers-back-up-our-obsession-with-phones – last access 7.5.14

[2] http://www.nielsen.com/us/en/newswire/2014/how-smartphones-are-changing-consumers-daily-routines-around-the-globe.html – last access 26.2.14.

QR code with their device and opening the contained link in their phone's browser. The completion code was only displayed, if the HTTP user agent string matched a known mobile browser. Additionally, we included several attention check questions throughout the survey.

The survey consisted of four main parts. First, participants were asked about their smartphone use in general, including why they do or do not use a code to lock their phone and which lock screen they use. In the second part, we captured how participants value their smartphone and which risks they consider when reasoning about their phone's security. Next, we asked participants about extra measures they take to protect their phone and in which situations they take them. In the third part, participants were asked whether or not they previously had security related incidents with their smartphone. If they indicated that someone previously had unwanted access to their smartphone, we invited them to report on the most severe case, using the critical incident technique [63]. In the last part, we collected demographics and IT experience. The questionnaire can be found in Appendix C.1.

We used open-ended questions to ask about extra measures, the reasons why participants do (not) lock their phone, as well as critical incidents. While there were too few critical incidents reported to justify coding, we coded the reasons and extra measures using an inductive coding approach. Two of the authors independently went through the answers and created codes. To capture as many facets of participants' answers as possible, codes did not represent complete responses, but certain common aspects, such as protection goals or likely attackers. The codeplans were then discussed and merged before both authors coded all responses, assigning multiple codes to each response. Conflicting codings were again discussed and resolved before a third coder independently coded all responses again using an improved codeplan. The final round of coding yielded no more conflicts. The final codeplan can be found in Appendix C.2.

5.3.2 Participants and Results

After pretesting the survey in the lab and on MTurk, 320 workers accepted the task in November 2013. We removed 60 response sets due to incorrect completion codes (i. e. the smartphone check failed), implausible timing, or wrong answers to two or more attention check questions. Participant demographics are summarized in Table 5.1. Participants indicated high IT expertise as almost a quarter has worked in or studied IT and 39.6 % reported the highest value when asked how they rate their own understanding of computers and the Internet. All participants indicated that they use their smartphones on a daily basis with the majority using them at least once per hour. Mobile operating systems were evenly split between iOS and Android. Interestingly, 51.2 % of participants indicated that they have suffered from a smartphone related incident before.

Overall, 42.7 % of participants indicated that they use some form of lock screen, including a PIN, a password or an unlock pattern, but not including the "slide-to-unlock" mechanism. In the remainder of the paper, this will be referred to as

N	260	
Age	18 – 67 years	
	median 31 years	
Gender	45.4 %	female
	54.6 %	male
Occupation	50.8 %	full-time employee
	13.1 %	part-time workers
	10.0 %	self-employed
	9.2 %	student
	7.3 %	unemployed
	9.6 %	other
IT Experience	22.7 %	have worked in or studied IT
IT Expertise	39.6 %	very high self-rating
Smartphone Use	36	months (median)
Usage Frequency	79.2 %	hourly or more often
Mobile OS	49.0 %	iOS
	48.7 %	Android
	2.3 %	Other
Lock Screen	40.9 %	Slide-to-Unlock
	33.6 %	PIN
	8.5 %	Pattern
	0.8 %	Password
	16.2 %	None
Incidents	21.5 %	phone lost
	11.9 %	unwanted access
	8.5 %	stolen
	28.5 %	broken phone, lost data

Table 5.1: Online study participant demographics.

"code-lock". Split by operating systems, 55.2 % of iOS users were significantly more likely to have a code-lock compared to only 30.4 % of Android users (Fisher's Exact Test (FET), $p < .001$). Of the 22 users with the Android pattern unlock, only 2 indicated that they had made the lines between the dots invisible.

Locking Behavior

We asked the 111 users that use a code-lock, how frequently they think they unlock their phone on an average day. Answers ranged from 1 to 100 with a median of 20 and a mean of 24.3 times. Our field study will show that many participants significantly underestimate their phone use. Additionally, we asked these 111 users to rate their sentiments towards locking on a 5-point scale. 64.9 % were not or mostly not concerned that someone might be shoulder-surfing their code entry. 25.5 % somewhat or fully agreed that they desire an easier way of unlocking their phone, while 69.4 % somewhat or fully agreed that unlocking their phone is easy. Yet, 46.8 % also somewhat or fully agreed that unlocking their phone can be annoying. At the same time, 95.5 % somewhat or fully agreed that they like the idea that their

phone is protected. These results already show a certain ambivalence towards the code-lock mechanism.

Code	Count
Protection Goal	**88**
– Controlling access to phone	32
– "Safety"/"Security"	25
– "Privacy"	15
– Protection in General	6
– Increasing difficulty of unwanted access	8
– Increasing time to recover/remote-lock phone	1
– Enable data encryption	1
Protect information	**75**
– Information in general	38
– *Private* information in general	14
– Emails/Messages	9
– Photos	4
– Other app-specific content	5
– Confidential (work) information	5
Protect from specific scenario	**62**
– Lost phone	27
– Stolen phone	20
– Unattended phone	8
– Pranks/someone "messing up" phone	5
– Misplaced phone	2
Protect from attacker	**55**
– Unspecific	32
– Unwanted person	11
– Own children	11
– Roommates	1
Other	**38**
– Protect certain action	17
– Mandatory lock screen	6
– Context (work/death)	4
– Other motivation	11

Table 5.2: Reasons for using a code-based locking mechanism. Bold counts are sums of sub-counts.

Locking Motivation

When asked why the 111 users with a code-lock chose this protection, answers were centered around four topics: protection goals, protection of information, protection in specific scenarios, and protection from attackers. An overview of the 318 code instances we tagged answers with can be found in Table 5.2. Participants provided a very diverse set of reasons across the four main topics. However, individual participants justified their choice using only few of the available aspects (ranging from

1 to 6 codes per participant, median of 1.0). While many answers were unspecific ("to protect my information"), other participants provided well reasoned answers, such as increasing the time an attacker needs to access the data. It is also noteworthy that no participant mentioned protecting login credentials or logged-in accounts directly.

We asked the 149 participants without a code-based lock why they chose not to have any protection mechanism for their phone. Table 5.3 provides an overview of the 236 code instances we attached to the answers. In this case, answers were mostly centered around two issues, namely inconvenience and the absence of a threat. Answers again included reasonable choices, such as choosing not to have a lock screen because the contained data is not considered sensitive by the respondent, while others were less rational, such as "I don't feel like putting a password on it".

Code	Count
Absence of threat	**118**
– don't need security	25
– nothing to hide	23
– no sensitive data	16
– keep phone physically secured	29
– use only in private environments	11
Inconvenience	**85**
– Too annoying	3
– Takes too much time	23
– Use phone too frequently	13
– Mental burden	3
Negligence/Carelessness	**8**
Dislike Locking	**7**
Other	**25**
– locking causes problems	12
– protect phone using another measure	6
– Other reason	7

Table 5.3: Reasons for not using a code-based locking mechanism. Bold counts are sums of sub-counts.

Smartphone Risks

To assess which risks to the content on their phones participants are most concerned about, we asked them to select the worst thing that could happen to their phone from a list of six statements (cf. Appendix C.1). 52.7 % stated that losing the phone itself is worst as they would have to buy a new one. This result shows that, for many users, the monetary value of the hardware is more important than the associated privacy and security risks for accounts and data. However, such risks were mentioned second-most: 20.0 % chose losing the data that is on the phone in general as the worst possible scenario, while 11.9 % chose account abuse on a lost phone and 8.8 % data abuse on a lost phone. Only 4.2 % and 1.2 % chose app

abuse and data abuse respectively on an unattended phone. It has to be noted that lock screens cannot protect devices from getting lost and data loss is usually more influenced by backup strategies than authentication mechanisms. Therefore, 26.1 % of these scenarios could probably be prevented using adequate security mechanisms. The remaining 1.2 % of participants stated a combination of these six scenarios or gave another scenario. While the figures only relate to risks participants were most concerned about, these also likely influence users' behavior most.

Participants were also asked to to rate each of the six worst case smartphone risk scenarios in terms of severity and likelihood, the two classic dimensions applied to evaluate risk. We also included a third dimension, presence, that measures how frequently this risk is on a participant's mind. While the first two dimensions can capture a "value" of this risk, the third attempts to quantify how much this value influences day-to-day decision making. A risk that is considered very important by users is not only one that is particularly severe and likely but also one that is present in the users' minds frequently. In terms of presence, all six risks were on users' minds similarly infrequently: for all six risk scenarios, 65 to 82 % of participants indicated that they think of this risk infrequently or very infrequently. A Friedman's ANOVA across the six scenarios did not yield a significant difference ($\chi^2(5) = 7.74$, $p = .17$). Similarly, the likelihood of the six scenarios happening to oneself was rated as likely or very likely only by 14 to 21 % of participants. Again, these values were not significantly different ($\chi^2(5) = 1.96$, $p = .85$). There was, however, a highly significantly different rating of risks in terms of severity ($\chi^2(5) = 62.17$, $p < .001$): Post-hoc tests with Bonferroni correction revealed that losing the phone and having to replace it was considered more severe than losing data or having unwanted access to the phone. In addition, participants believed that risks to data and accounts are more severe when a phone is lost compared to when the phone is unattended.

We also asked participants to compare their individual smartphone worst case to other negative situations in other contexts on a 5-point numerical scale from "not as bad" to "similar" to "worse". The situations comprised losing data on their PC, losing their wallet, losing the key to their home or their car, getting their email account hacked or someone breaking into their home. Someone breaking into one's home was rated as somewhat worse or worse by a majority of 86.5 %. Losing the key to their home or car was rated as not as bad or similar to the worst case smartphone scenario by 60.0 % and 47.7 % respectively. Also, losing data on their PC was rated as not as bad or similar by 56.2 %. Getting their email account hacked or losing their wallet ranged in between someone breaking in and the three other scenarios. This indicates that users may be ready to invest as much effort into protecting their phones as they are to protect themselves from losing the key to their home or data on their PC.

We then asked participants to rate which kinds of attackers are most likely to attempt unwanted access to their smartphones. They rated four potential attackers, known malicious and known curious as well as unknown malicious and unknown curious, on a 5-point scale from very unlikely to very likely. We found a highly significant difference between the four attackers (Friedman's ANOVA, $\chi^2(3) = 40.07$, $p <$

.001) and Bonferroni-corrected post-hoc tests showed that the known curious and the unknown malicious attackers where considered more likely than the two other attackers.

For those participants who rated a known attacker as neutral, likely or very likely, we also asked whether or not they considered eight types of known persons as a potentially curious or malicious person for their rating. The most frequently chosen types of persons are outlined in Table 5.4.

Curious Attackers		**Malicious Attackers**	
Attacker	**Freq.**	**Attacker**	**Freq.**
Close Friends	73.2 %	Other known people	68.9 %
Acquaintances	54.3 %	Co-workers	29.3 %
Parents	53.0 %	Acquaintances	25.0 %
Children	51.8 %	Friends of friends	23.2 %
Friends of Friends	46.3 %		

Table 5.4: Kinds of persons respondents considered as known malicious or curious attackers.

Extra Measures

To see how participants cope with risks to their smartphone besides inbuilt protection measures, we asked them if they sometimes apply extra measures to protect their phone. 83.5 % indicate that they keep the phone on their person or in their bag, 50.8 % leave the phone in a safe place and 33.5 % enable a lock screen or choose a harder unlock code for certain situations. Furthermore, we asked if participants with code-lock screens take some of five measures against shoulder surfing: 27.7 % indicated that they tilt their screen away while entering their unlock code when shoulder surfing is possible, 16.2 % wait a moment, 11.2 % turn around, 8.8 % cover phone, and only 7.3 % have previously changed their unlock code after a potential shoulder surfing happened. We also prompted participants to give up to three situations in which they apply those measures. As participants often not only listed a situation but also additional measures, we coded these responses for both concepts. We attached 701 instances of situation codes and 248 instances of measure codes, while each answer could receive multiple measure and situation codes. The corresponding codeplans can be found in Appendix C.2.3 and C.2.4.

In addition to the protection measures we already asked about, the coded responses revealed that in 45 instances participants mentioned to be paying extra attention to their phone. In 19 instances, other technical measures, such as turning the phone off, encrypting data, relying on remote wiping and locking functionality, removing the memory card or having a backup were quoted. With respect to situations, we found that most participants referred to public or semi-public spaces as situations where they would need extra protection. Examples include being “out” in general (59), going to events or concerts (23), while being at a gym or during workout (42), during parties or in bars (35) or at work (52). A feeling of unfamiliarity or

unknown spaces were mentioned in 50 instances as were discomforting spaces, such as dark areas or dangerous neighborhoods (24). However, private spaces, such as a home, were also perceived as situations were extra measures may be necessary (16). Leaving the phone in the car (21) or uncontrolled situations where a phone is left unattended or one is less cautious (102) were frequently mentioned. In addition to unspecific unattended situations (71), participants mentioned leaving the phone to charge, while sleeping or drinking or when bags are handed over for example at the airport. Persons were also often a component of situations that were protected with extra measures (overall 61 instances): unfamiliar or untrusted persons (20), other people in general(15), kids (9), (ex-) partners (4), friends (6), and coworkers (2) were all mentioned. Finally, device sharing (5) or having sensitive and inappropriate data (4) were also quoted as situations were extra measures need to be taken.

Critical Incidents

The 31 participants who reported having been victim of unwanted access before, quoted the following critical incidents during which unwanted access happened: children or siblings accessing the phone for fun, snooping (ex-)partners, friends playing pranks and abusing accounts, a thief acquired the phone, friends snooping on private information, a stolen phone that was sold and then returned to the police by the buyer because the phone was not wiped, parents "checking" on their children, and having a virus on the device. We then explicitly asked about the harm that arose in this situation: ten participants stated an invasion of privacy, four got into a conflict with the other person, accounts were abused in three cases, others were offended in three cases and embarrassment was caused in one case. Seven participants reported that they were frustrated or mad and six participants indicated that they saw no harm in this incident. On the other hand, we asked participants what good came from the incident. Responses included clarified relationships and boundaries in five cases, a new phone in one case, five participants stated to have learned to pay more attention to their phone (even though they are still not locking their phone) and one started using a lock mechanism. For eight participants, nothing good came from the incident. In terms of having a code-lock or not, these critical incidents show that many of them could have been prevented by using a code-lock. However, as the previous subsections have shown, a large number of reasons let users choose not to have a code-lock. Also, the intangible nature of the consequences may also play a role in the adoption decision. The next chapter will discuss this issue in more detail.

Summary

The online survey showed that security and privacy problems of smartphones are concerns for participants. However, they are also often not top concerns in comparison with other day-to-day risks. The smartphone hardware itself appears to be more valuable to users than the security and privacy risks for data and accounts. Users locking their phones state a large variety of reasons for this behavior, which is

often based on unspecific and abstract needs and threats. Similarly, reasons for not locking the phone were mostly based on convenience or the absence of a perceived threat. Participants believed that most risks occur in public or semi-public spaces and arise from unknown malicious or known curious attackers. We also collected a plethora of additional measures users take to protect their phones beyond lock screens. Participants also referred to these to justify not having a lock screen and we believe that physical precautions are a major influence to this view. Overall, the survey shows that locking one's phone is not a primary concern and the risk of shoulder surfing is not present in many users' minds. The desire for protection also differs largely by context.

5.4 Longitudinal Field Study

While the survey results already provide interesting insights, they are based on self-reports at one point in time. To further evaluate the role of context to unlocking and hence generate ground truth for improvements of smartphone locking schemes, we conducted a longitudinal field study with 57 participants over four weeks. The design of the study was governed by three research questions: How frequently do people unlock their phone? What is the influence of context on perceived necessity of locking? And how frequently are users potentially subject to shoulder surfing or unwanted access to their device?

To increase data validity, we instrumented users' private phones to implement an experience sampling method and gather quantitative data like unlock frequencies and authentication times. The field study was grounded on the results of the online survey and a focus group. We conducted this focus group (n=7) to familiarize ourselves with participants' reasoning and views on our research questions. The results helped us to further reduce the question list to the most important aspects and keep the participants' additional effort as low as possible.

5.4.1 Method

To elicit a longitudinal picture of users' everyday behavior and perceptions, a subtle and low-effort data collection method was necessary. We decided to collect data from users of the Android OS, as it is both very common and provides suitable APIs to collect the desired data. Andreas Fichtner implemented an app that would automatically log (un)locking activity on users' phones. Additionally, the app displayed mini-questionnaires on random occasions to obtain a sample of users' views on their locking behavior immediately and within a given situation (cf. below for details). The logged information was periodically backed up to our servers when the phone was connected to a WiFi network. We collected data over a period of four weeks.

Presenting questionnaires in-situ is known as the experience sampling method (ESM) and has been previously applied to investigate real-life situations [92, 33, 87]. Other longitudinal methods have been used to capture user experience on mobile phones

[107], such as the Day Reconstruction Method. However, for our exploration of smartphone locking behavior and perception, we can easily use the capabilities of modern smartphones to collect the necessary data in situ and do not need to let users remember parts of their experience. Additionally, Möller et al. have previously demonstrated problems with relying on self-reporting during long-term studies [122]. We hence split our data collection efforts into two parts: Activity Logging and Mini-Questionnaires. The questionnaires only ask for immediately observable information or information from the near past and thus do not need to heavily rely on participants' memory. We present the details of our approach in the following two subsections.

Activity Logging

Our app monitored SCREEN_ON and SCREEN_OFF intents as well as the KeyguardManager state provided by the Android OS. This allows us to derive when a device was activated, unlocked and deactivated. Figure 5.1 provides a state-machine representation of the collected state information. Whenever a users presses the hardware button activating or deactivating the smartphone's screen, the state transitions between ON_* and OFF_* states. When the lock screen is dismissed, the system transitions from ON_LOCKED to ON_UNLOCKED. Finally, the transition from *_UNLOCKED to *_LOCKED occurs either immediately or after a certain delay, depending on users' configurations. We logged timestamps when entering a state. It is important to note that especially the time it takes to unlock the phone (transitioning from ON_LOCKED to ON_UNLOCKED) is a worst-case estimate, as it includes the time users spent viewing notifications or the clock on the lock screen first. Also, our app did not need any permissions to collect this data.

Mini-Questionnaires

As we aimed to capture participants' perceptions of threats related to their smartphone locking behavior in their daily life, we enriched the automatically logged data with participants' subjective views. We hence applied a method somewhat similar to what Cherubini and Oliver proposed [30]. Using two very short questionnaires, participants were asked about their surroundings and subjective perceptions. The two questionnaires were randomly displayed with a certain probability after a subset of device unlocks and contained multiple-choice questions to facilitate rapid answering. One questionnaire focussed on the unlock procedure itself and gathered shoulder surfing possibilities, who an attacker would be, as well as how likely and severe such an attack would be. Participants were instructed to briefly consider their environment and indicate if someone was able to see the contents of their screen in this situation. Additionally, we elicited satisfaction with the locking procedure in this situation and the sensitivity of the data to be accessed. Participants were instructed to judge sensitivity of data subjectively without giving them any further definition in order to not disrupt their own mental model. The second question-

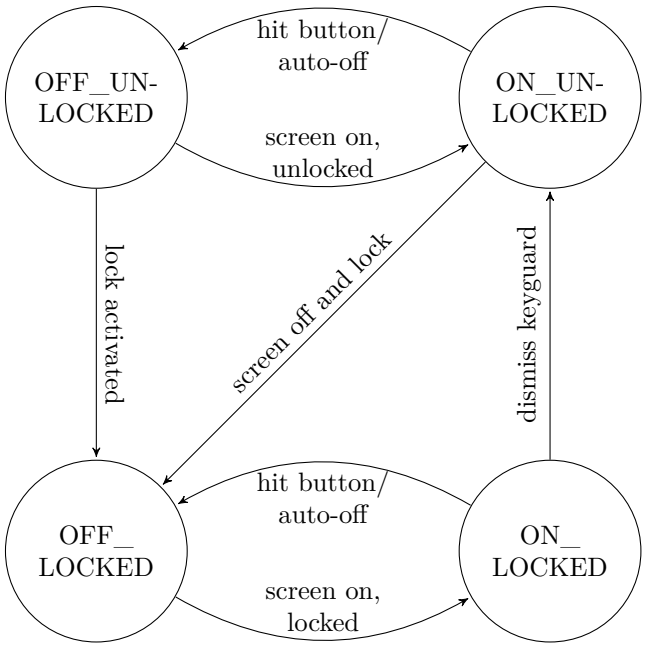


Figure 5.1: The states and transitions logged during data collection. ON and OFF parts of the node labels derive from the screen state, (UN)LOCKED denotes the keyguard state.

naire focussed on the time span between the current unlock and the last use. This questionnaire elicited views on the necessity of the lock screen, if unwanted access has been possible, and how annoying the locking mechanism was in this situation. Both questionnaires asked participants to characterize the environment they are currently in as private, semi-public or public, according to the categories we obtained in the online survey as well as the pre-study focus group. The contents of both questionnaires can be found in the Appendix C.3.

Situation Sampling

To obtain a representative sample of day-to-day situations, we needed to randomly choose unlock events throughout the day after which we would display one of the two mini-questionnaires. Pre-testing showed that unlocking behavior varies widely between participants, days, and time of day. We hence dismissed the possibility to apply a fixed sampling schedule for all participants. Some participants may use their device more frequently during the day, while others may become particularly active in the evening. Additionally, we aimed to sample as many different situations as possible and therefore did not want to restrict the sampling time frame to, for instance, working hours as has been previously done in similar contexts [87]. Pre-testing also revealed that it takes about 30 to 40 seconds to complete the mini-questionnaires on the device. In order to not overwhelm participants, one of the two questionnaires would be randomly displayed with a certain probability and at most once per hour. Participants were also able to press a “Not Now” button, that would

dismiss this questionnaire immediately, in order to allow quick access to the phone if necessary.

At deployment time, the probability that a questionnaire was shown for a given unlock was set to 20 % based on the results of a one week pre-study. After one week of data collection, the probabilities were adjusted to collect about 5 to 6 questionnaires per day to keep the task as unobtrusive as possible while collecting a wide range of situations. Heavy users (at least 9 unlocks per hour) were throttled to 10 % selection probability and medium users (between 4 and 8 unlocks per hour) to 15 %. We chose to adapt the sampling rates to put an even burden on all participants and make the study less intrusive for participants that use their phone more frequently.

Briefing and Debriefing

All participants were briefed about the study and the method during an initial meeting in person or by phone. The data collection procedure and the questions in both questionnaires were explained and participants had a chance to ask questions. The app was then installed on each participant's phone before participants tested both mini-questionnaires. After the data collection period, participants came in for a debriefing interview. We collected the data from participants' phones and removed all traces of the app.

We also conducted a short interview, eliciting how participants liked the data collection method, whether any problems occurred, and if answering the questionnaires actually changed their behavior. We then presented a personalized data sheet to every participant, summarizing their phone use as well as survey answers. Participants' initial reactions to this data was recorded. We asked them whether they thought that the data adequately represents their everyday use, and whether or not we may have missed shoulder-surfing or unwanted access occasions. We also asked them questions on their attitudes towards code locks and smartphone use and collected demographics. The results of the debriefing sessions will be detailed in Section 5.4.2.

Participants

We recruited 57 participants at two locations in Germany in January 2014. At one location, 27 participants were recruited through message boards, social networks, and mailing lists, while at the other 30 students and graduates where recruited using a study participation mailing list. We advertised a four week study on Android lock screens for users that have had a smartphone with Android 2.3 or higher for at least 3 months. A 10 Euro base-salary plus 14 Euro-cent per completed mini-questionnaire were promised as compensation. Participants earned 30.79 Euros on average.

While all 57 participants completed the data collection part of the study, we removed one participant who did not show up for debriefing, three participants who repeatedly modified the time on their phone during data collection, and one participant

where data collection failed for several days, as our app did not restart after rebooting this user's device. The remaining 52 participants' demographics are summarized in Table 5.5. While our participants mainly comprise students of which about half also have some IT experience, we find that this is a population worth studying as they are often very active and experiencing a wide range of situations, but also have phases where they sit in front of a desk for extended periods. As we aim to explore how different environments influence locking behavior and risk perception, our sample offers a good chance to collect a wide range of usage contexts. However, it still has to be noted that our results cannot be generalized to any particular population.

N	52	
Age	19 – 32 years, median 23 years	
Gender	23	female
	29	male
Occupation	47	undergrad or grad students
	5	PhD student or staff
Highest degree	34	high school diploma or less
	18	Bachelor/Master degree
IT experience	25	work(ed) in or study(ed) IT
Smartphone history	34	months (mean)
Lock screen type	13	PIN
	22	Pattern
	17	Slide-to-unlock
Code lock for	22	months (mean)
Avg. PIN length	4.5	digits (range: 4-6)
Avg. Pattern length	5.2	cells (range: 4-8)

Table 5.5: Longitudinal field study participant demographics.

5.4.2 Results

Participants contributed 29.5 days of data on average. To equalize the time we analyze per user, we pruned each participants' dataset to 27 complete days from midnight to midnight by removing the first hours and the remaining days. Due to our method, each user contributed a different amount of data. In order to not over-represent users that use their phone more frequently, we first aggregate data per user and then average across users' aggregates where appropriate.

Logged Data

Within the 27 days, we observed an average of 2242.3 activations (switching the screen of the device on) per participant ($sd = 1160.2$, median=2260), ranging from 651 to 5419. Correspondingly, 1286.0 unlocks (dismissing the lock screen after activating the phone) were logged on average per participant ($sd = 711.8$, median=1127), ranging from 215 to 3545.

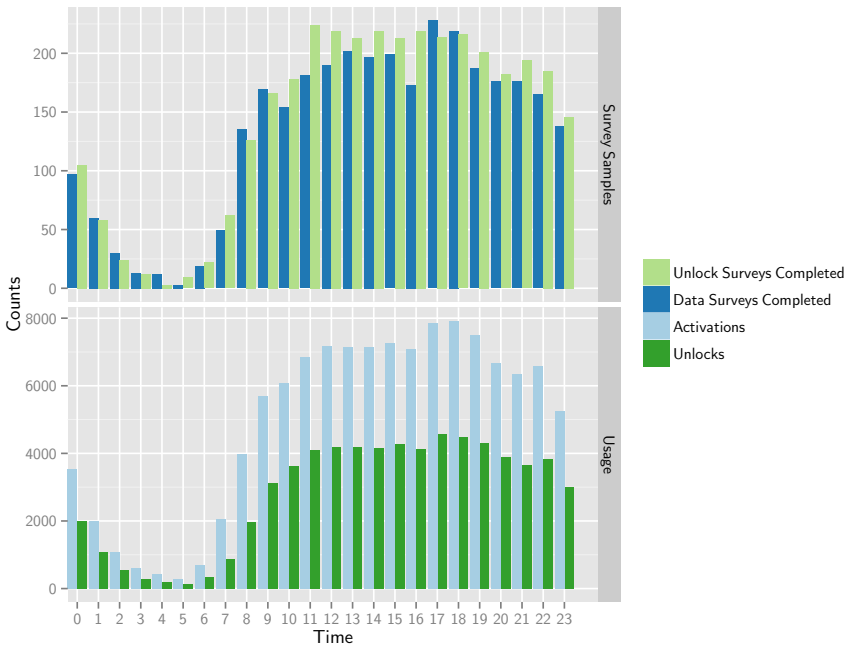


Figure 5.2: Overview of sampled situations per time of day, comprising number of mini-questionnaires shown as well as cumulative number of activations and unlocks.

Per day, participants activated their phone 83.3 times (sd = 43.0, median=83.8) and unlocked 47.8 times (sd = 26.4, median=42.1) on average. This translates to an average of 5.2 activations and 3.0 unlocks per hour, assuming that a user is awake for 16 hours per day. Participants unanimously attributed the discrepancy between activations and unlocks to activating the screen of their phone to see the current time and to check for notifications. Overall, usage was largely similar during daytime hours, ramping up in the morning and down in the evening after 9 pm. The histograms in Figure 5.2 provide an overview of all participants' aggregated use (bottom facet) by time of day during the experiment (27 days). The top facet shows the corresponding number of mini-questionnaires of both types participants completed.

During recruitment, we asked participants how frequently they think they unlock their phone per day. Figure 5.3 compares these guesses with the measured frequency. We find that most users severely underestimated their use. However, participants who use their phone less frequently appeared to give better estimates.

Figure 5.4 shows that the distribution of unlocks per hour across users is bimodal. We hence group users into heavy and "regular" users, where heavy users unlock their phone more than 3 times per hour. Please note that significance testing results based on this grouping are only of exploratory nature, as groups were formed post-hoc.

Activating and unlocking the phone took 2.67 seconds without a code lock (sd = 8.46s, median=1.26s), 3.0 seconds using a lock pattern (sd = 13.3 sec, median=1.69s),

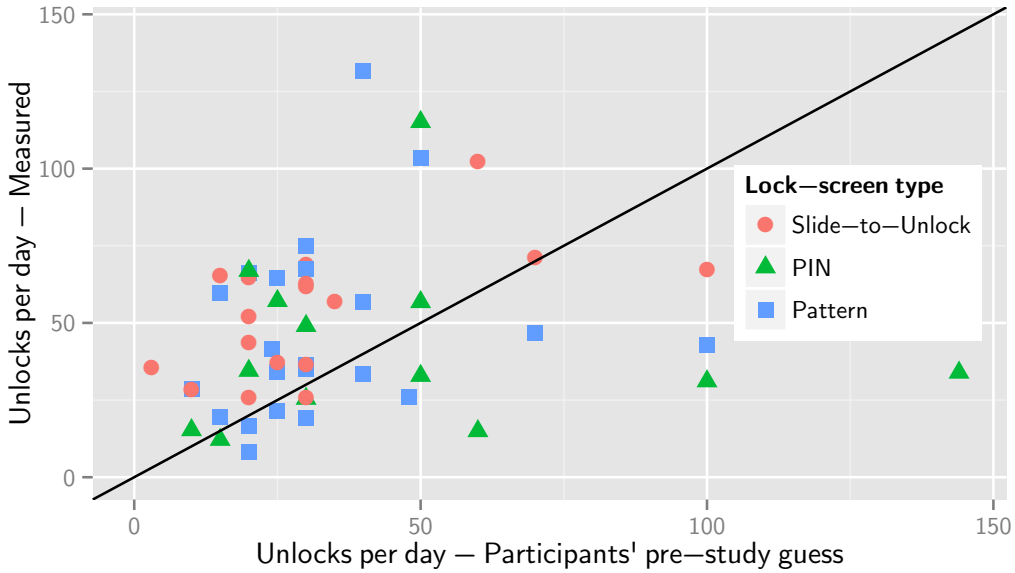


Figure 5.3: A comparison of participants' pre-study guess of unlocks per day versus the actually measured values. The closer the points are to the diagonal line, the more accurate their guess was. Most users underestimated the frequency of their unlocks.

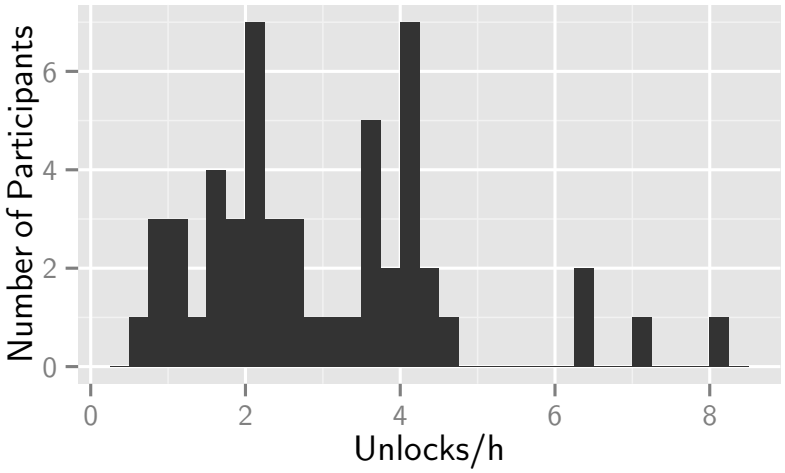


Figure 5.4: Histogram of users' mean combined activation and unlock times.

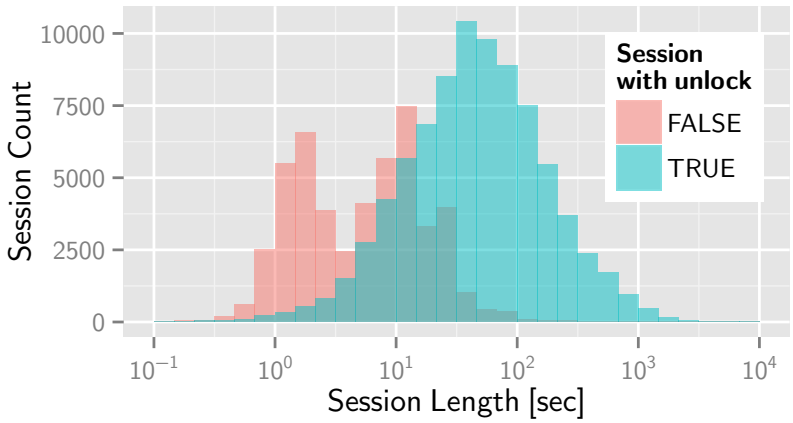


Figure 5.5: Histogram of session lengths on a log scale.

and 4.7 seconds using a numeric PIN (sd = 20.72s, median=2.85s) across all unlocks. Averaging unlock times per user, we ran a user-type by lock-type between-subjects ANOVA and found a highly significant main effect for lock-type ($F(2, 46) = 11.37$, $p < .001$) as well as s significant main effect for user-type ($F(1, 46) = 6.39$, $p = .002$). Heavy users completed their unlocks more quickly on average (2.9 vs. 3.8 seconds). Holm-corrected pairwise testing also showed that PIN (4.9 seconds on average) was significantly slower than the two other mechanisms (Slide-to-Unlock 2.6 and Pattern 3.2 seconds, $p < .001$, Cohen's $d = 1.58$ and 1.27 respectively). During the 27 days of the experiment, participants spent an average of 1.17 hours each ($sd = .87$, ranging from .2 to 5.1 hours) just unlocking their device.

An average session (from SCREEN_ON to SCREEN_OFF) lasted 70.3 seconds ($sd = 241.5s$). However, sessions where participants actually saw the home screen lasted for 104.1 seconds (sd = 193.9s, median=45.6s) on average, including the time it took to dismiss the lock screen. The remaining sessions (those when the device was not unlocked) lasted only 12.4 seconds (sd = 297.6s, median=5.2s) on average. Figure 5.5 gives an overview of session lengths, grouped by whether or not the session entered the home screen. It can be clearly seen that sessions last longer once the lock screen was dismissed. Also, the distribution of session lengths on a locked device is bimodal. We hypothesize that the maximum at about one second is for checking the time, while the maximum at about 10 seconds session lengths represents cases where users check notifications.

Averaging per user, heavily unlocking users spent highly significantly shorter sessions on average (40.5 seconds) than regular users (50.9 seconds, Welch's t-test, $t = 16.05$. $p < .001$). Overall, users spent 43.0 hours on average (sd = 22.1h, median=41.2h) using their smartphone within the 27 days of our experiment, of which an average of 2.9 hours were spent on a locked device (i. e. checking time or notifications on the lock screen). 2.9 % of the overall time was related to unlocking the phone on average, ranging from .6 to 9 %.

Mini-Questionnaire Data

We collected 3410 completed unlock risk questionnaires (65.6 per user on average, range 15-110) and 3172 completed data risk questionnaires (61.0 per user on average, range 15-105) with the experience sampling part of our study. The sampled situations included a wide range of times of day and even collected samples when participants used their phone at night (cf. Figure 5.2 above). Filling the questionnaires took 23.7 ($sd = 35.9$) and 21.3 ($sd = 22.6$) seconds on average for each type respectively. In the following, we present results from questionnaire parts individually.

Environments In both questionnaires, participants reported the environments in which they were in the moment they unlocked the phone or in which they have been since they last used the phone. Averaging environment proportions per user, these environments were mostly private (62.4 %), semi-public in 19.5 % of cases and public in 18.2 %. In line with previous findings [80], this indicates that most smartphone use takes place at home or in similarly private spaces.

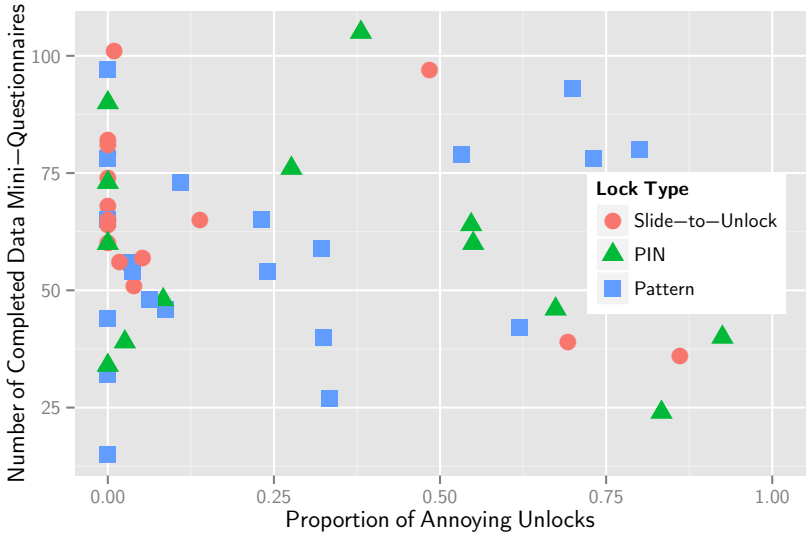


Figure 5.6: Proportion of annoying unlocks per user versus how many questionnaires were completed.

Perception of Lock Screen In the first mini questionnaire, we asked participants how annoying the unlock (which they just completed prior to filling out the questionnaire) was. Participants reported different proportions of annoying unlocks (either "annoying" or "very annoying"). Figure 5.6 shows the relationship between the proportion of annoying unlocks, the number of completed questionnaires (corresponding to how heavily users use their smartphone), and the type of lock screen they use. A large amount of participants was very happy with their lock screen, as they re-

ported no or almost no annoying unlocks across their questionnaires. Only 12 of 52 participants indicated being annoyed by their lock screen in more than 50 % of their mini-questionnaires. There also is no clear trend of users with a particular lock type being more annoyed. However, we note that only three users with Slide-to-Unlock reported annoying unlocks in more than a quarter of their questionnaires.

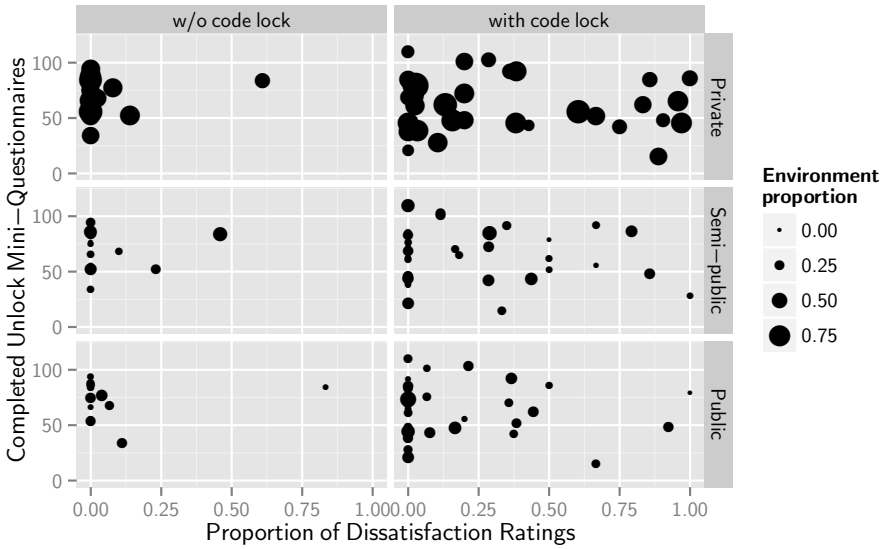


Figure 5.7: Proportion of dissatisfied ratings per user versus how many questionnaires they completed, grouped by whether they had a code lock screen and in which environment the rating was given. Point size indicates the relative frequency with which this user reported being in each environment.

Additionally, in the other mini questionnaire, we asked if users with a code lock would have rather not had a code lock in this situation and vice versa. High ratings on the 5-point numeric scale of this question indicate dissatisfaction with having a code lock or not. Figure 5.7 and Table 5.6 give an overview of the answers provided. In the figure, the y-axes additionally show how many questionnaires each user completed, approximating how frequently the phone is used. Participants' answers are grouped by the environment they were provided in and whether or not this participant had a code-lock. The size of each point in the graph indicates how frequently this user reported being in this environment.

The data shows that participants without a code lock were generally more satisfied with their status quo. Only few of them indicated dissatisfaction in more than a quarter of their responses across all environments. Participants with code locks showed more variability and more participants indicated dissatisfaction in more than a quarter of their responses. Especially users that are frequently in private environments were more frequently dissatisfied with their code locks. It is also noteworthy that fewer code-lock participants indicated strong dissatisfaction in public environments compared to semi-public or private situations.

Environment	# *Situations*	Mean Proportion of Dissatisfaction Ratings w/o code lock	with code lock	overall
private	2115 (62.0 %)	5.0 % (14.9 %)	32.7 % (36.0 %)	23.6 % (33.2 %)
semi-public	690 (20.2 %)	4.6 % (12.2 %)	23.0 % (29.3 %)	17.0 % (26.3 %)
public	605 (17.7 %)	6.2 % (20.1 %)	16.6 % (26.9 %)	13.2 % (25.2 %)
Overall	3410	5.3 % (15.8 %)	24.1 % (31.4 %)	17.9 % (28.6 %)

Table 5.6: Participants' dissatisfaction with their locking mechanisms by environment. Numbers in brackets in the last three columns indicate standard deviations.

A possible interpretation is that being annoyed by a lock mechanism overlays risk perception to some extent as there is only a limited trend towards more satisfaction with lock screens in potentially more dangerous public situations.

Data Sensitivity We asked each participant for subjective ratings on how sensitive the data that is going to be accessed in this session is. In 684 (20.1 %) of 3410 completed mini questionnaires, users indicated that they did not know what kind of data they were going to access. Aggregating proportions of unknown accesses per participant, the mean proportion amounts to 19.6 % ($sd = 25.0$ %) and participants' individual values range from 0 % to 88.2 %.

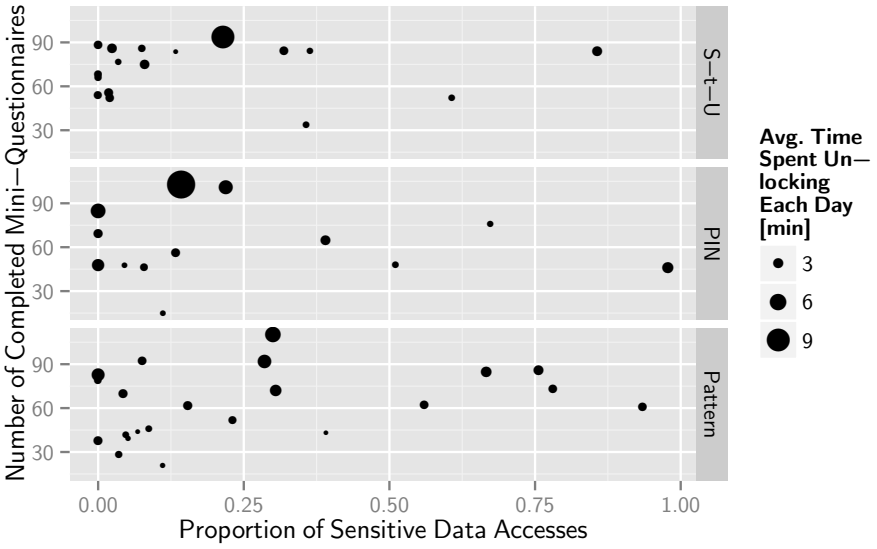


Figure 5.8: Proportion of sensitive data accesses per user, grouped by lock mechanism. The y-axis also shows completed questionnaires and the size of the points indicates how many minutes a participant spent every day activating and unlocking his or her phone on average.

In 25.3 % (691) of the 2726 remaining reported situations, participants indicated accesses to sensitive data. For each user, this means that during the experiment only 10.6 hours ($sd = 15.0$) of the 43 hours each participant spent using their device

contained accesses to sensitive data on average. All but ten users indicated that they access less sensitive data in more than half of the sampled sessions. Figure 5.8 visualizes the proportions of sensitivity ratings across the sampled situations per participant. It is also visible that one user spent a lot of time unlocking the phone each day even though the data that should be accessed was not sensitive in most cases. Notably, the ten participants that were accessing most sensitive data use their phone more frequently (i. e. filled more questionnaires).

Environment	*# Situations*	**Known Person**	**Unknown Person**	**Nobody**
private	2115 (62.0 %)	8.6 % (181)	0.0 % (1)	91.4 % (1933)
semi-public	690 (20.2 %)	22.2 % (153)	4.6 % (32)	73.2 % (505)
public	605 (17.7 %)	10.4 % (63)	24.5 % (148)	65.1 % (394)
Overall	3410	11.6 % (397)	5.3 % (181)	83.0 % (2832)

Table 5.7: Shoulder surfing possibilities across potential "attackers" and environments.

Environment	*# Situations*	**Unlikely**	**Low Severity**
private	2115 (62.0 %)	56.6 % (103)	92.9 % (169)
semi-public	690 (20.2 %)	65.4 % (121)	84.9 % (157)
public	605 (17.7 %)	56.0 % (118)	68.3 % (144)
Overall	3410	59.2 % (342)	81.3 % (470)

Table 5.8: Percentages of user ratings providing low likelihood and severity ratings with respect to possible shoulder surfing attempts (i. e. by known or unknown persons).

Shoulder Surfing Table 5.7 gives an overview of shoulder surfing possibilities perceived by our participants. Across the 3410 unlock risk mini-questionnaires we collected, shoulder surfing was not perceived to be possible in a majority of 83.0 % of cases. When it was possible, mostly known persons were observers, except in public environments. In more than half of the situations where shoulder surfing would have been possible, participants thought it to be unlikely or very unlikely that this did actually happen (cf. Table 5.8). Had it happened, the threat from the potential attacker would have been low or very low in most of the possible shoulder surfing situations, especially in private environments. Overall, we found only 11 of the 3410 (.3 %) reported situations were it was likely that a shoulder surfer was looking at the screen and it would have been severe or very severe if that had actually taken place. Seven of these occurred in public situations.

We also asked those participants with a code lock whether or not they protected their code entry during the last unlock, by for example tilting their screen away from onlookers or waiting to unlock the phone. Only 18 participants reported 52 instances in which they actively protected the code input from a shoulder surfing threat within 1869 sampled situations where a code was entered (2.8 %).

Unwanted Access In the data risk mini-questionnaire, participants were asked to report situations in which unwanted access to their smartphone was possible. Eleven participants did not report any of these situations and the remaining 42 participants reported a total of 245 occasions out of 3172 possibilities (7.7 %) and between one and twenty occasions each. Table 5.9 provides an overview of unwanted access occasions and who an attacker would have been. Additionally, Table 5.10 details how many of these occasions were rated as unlikely and for how many the consequences participants saw were rated as benign. Unwanted access was infrequently possible, mostly by known persons except in public situations and rated as mostly unlikely and benign.

Environment	*# Situations*	**Known Person**	**Unknown Person**
private	131 (53.5 %)	97.7 % (128)	2.3 % (3)
semi-public	75 (30.6 %)	70.7 % (53)	29.3 % (22)
public	39 (15.9 %)	23.1 % (9)	76.9 % (30)
Overall	245 (7.7 %)	77.6 % (190)	22.4 % (55)

Table 5.9: Unwanted access occasions by environments and potential attackers.

Environment	*# Situations*	**Unlikely**	**Benign Cons.**
private	131 (53.5 %)	92.4 % (121)	86.3 % (113)
semi-public	75 (30.6 %)	93.4 % (70)	64.0 % (48)
public	39 (15.9 %)	79.5 % (31)	18.2 % (11)
Overall	245 (7.7 %)	90.6 % (222)	70.2 % (172)

Table 5.10: Percentages of users providing low likelihood and severity ratings of potential consequences with respect to reported unwanted access occasions.

Debriefing Interview

During the debriefing sessions, we asked participants if they had any problems with our data collection app. Participants reported no severe problems influencing their use of the phone. Minor issues included mini-questionnaires popping up when taking the phone away from the ear after finishing a call in some instances or questionnaires disappearing because of popups from other apps for a small number of participants.

We also asked if participating in the study or seeing the questionnaires influenced participants' smartphone use. One participant reported to have increased the time interval after which the lock screen is shown again from 30 to 90 seconds, another participant stated that he sometimes did not turn his screen off immediately. Three participants stated that they may have used the device a little less frequently at the beginning of the study. Ten participants said that being part of the study made them pay more attention to why and how often they use their phone. While it made them realize their usage, they reported not to have altered their behavior. One

participant said that he may remove his code-lock after the study, as participating made him realize how much effort unlocking with a PIN takes.

Participants were also asked to rate how annoying they found answering the mini-questionnaires to be. Only 5 participants selected 4 on a numeric scale from not annoying at all (1) to very annoying (5). 43 participants chose 2 or 3 and an additional 4 chose not annoying at all. On the contrary, many users reported that they found participating in the study very interesting for themselves, as it helped them assess their own behavior better. We also presented participants with a summary of the data they had shared with us, including frequencies of logged events, general usage statistics as well as overviews of mini-questionnaire answers. Most participants found these figures to be interesting and sometimes alarming, as they would not have expected to activate or unlock their phone as frequently. We also gave participants a numeric scale asking how well the collected data represents their actual behavior (logged data) and perception (questionnaire answers) from "not at all" (1) to "very much" (5). Participants felt that the data was valid: only one participant chose 3, 31 chose 4, and 20 chose 5.

To see how well the sampled situations covered participants' daily lives, we asked them if there were additional situations in which unwanted access was possible and if so of which nature those were. Several participants said that there probably were more of these situations, but they were mostly the same as the ones they reported in the sampled situations. Similarly, we asked participants if the proportion of situations where shoulder surfing was possible matched their own perception. Participants agreed that the numbers we collected and the proportion of shoulder surfing situations match their perception beyond the situations were questionnaires were shown. However, several participants mentioned that there were brief situations mostly in public environments where shoulder surfing would have been possible but no questionnaire was shown.

As in the online survey, we asked participants about previous critical incidents with their smartphone. Four participants had lost their smartphone before and two had unwanted access. In all cases, a lock screen was helpful to prevent more damage or was activated after the incident.

We also asked participants why they chose to have a lock mechanism with a code and coded results using the codes from the online survey. The 37 participants' answers contributed 115 code instances summarized in Table 5.11. Again, each answer was assigned multiple codes if several aspects were mentioned.

The results are similar to the online survey. However, in contrast to the online survey, several participants also gave restricting statements, noting that they do not believe that lock screens offer perfect security (7), that they do not really need security (6), or that others know their code anyway (3).

Again, participants without code locks also justified their choice and 15 participants contributed 44 code instances. Table 5.12 provides an overview of the reasons. The most frequently cited reasons for not using a lock, as in the online survey, are inconvenience and not seeing a threat.

Code	Count
Specific protection goal	13
Unspecific protection goal (“Security”)	12
Specific attacker	13
Unspecific attacker	10
Protect from specific scenarios (e.g. lost, stolen)	20
Protecting specific information	5
Protecting unspecific information	8
Protect from accidental input	4
Custom certificate	5

Table 5.11: Reasons for using a code-based locking mechanism of field study participants.

Code	Count
Inconvenience	17
Absence of threat	16
Locking causes problems	6
Protect phone using another measures	4
Not secure anyway	2

Table 5.12: Reasons for not using a code-based locking mechanism of field study participants.

Finally, we asked how sensitive participants consider the data on their smartphones to be in general and whether or not they share their code with other people. 22 participants (42.3 %) chose sensitive or very sensitive on a 5-point scale, while 23.5 % of users without a code-lock and 48.6 % of users with such a mechanism considered the data on their smartphones to be sensitive. However, this difference is only almost statistically significant (FET, $p = .076$). Unlock codes were shared with at least one person by 28 of 35 participants with a code-lock. Six participants indicated that at least 5 other people know their code. This also indicates that code-based locking mechanisms can be problematic in device sharing situations, as already noted by Karlson et al. [98].

5.5 Discussion

In the two previous sections, we presented results from two studies, which we summarize and discuss grouped by the most important observations in the following sections.

5.5.1 High Number of Unlocks

36 of 52 participants underestimated the number of smartphone unlocks by 141 % on average. This indicates that unlocking is a subliminal action in many cases and

unlock effort is kept low enough most of the time. However, even if a single unlock took only between 2.67 seconds (slide to unlock) and 4.7 seconds (PIN), the huge number of daily unlocks leads to a high impact of every additional second. Just over the course of our experiment, participants on average already spent about one hour unlocking their devices using traditional unlock screens. Taking into account that alternative authentication mechanisms often incur higher input times for increased security, this can easily add several hours of additional unlock time per month. This is especially critical when considering that average usage times per activation are relatively short and shows that authentication speed of feasible systems must be about as fast as PIN and patterns. Since our data indicates that unlocks are perceived as unnecessary in private environments and sensitive data is seldom accessed, we suggest that more effort should be put into researching how to decrease the number of unlocks by deploying usable context- and content-dependent locking mechanisms. The works of Hayashi et al. [80, 82] are a first step in this direction. Overall, increased effort in terms of time would certainly hinder adoption, especially since the often-cited protection against shoulder-surfing may not hold significant value for users.

5.5.2 Reasons For (Non-)Use Of Authentication Are Highly Diverse

The results of both studies suggest that reasons for using or not using protection mechanisms to access smartphones are highly diverse. Often, they are not based on objective reasons and were not valid from a technical perspective. In turn, a considerable number of participants provided reasonable justifications. Furthermore, others argue that code locking mechanisms are not perfectly secure anyway and even have drawbacks should the device be lost.[3] Participants without a code-lock in the field study were also very satisfied with their choice and indicated very few situations where they would have rather had a lock screen. In turn, dissatisfaction with a code-based lock was not as pronounced in public situations, as participants valued protection slightly more in that case. In terms of attackers, survey participants were most afraid of unknown malicious as well as known curious attackers. This is mirrored in the field study results, where known persons had the most shoulder surfing and unwanted access possibilities in private environments while unknown persons dominated in public situations.

5.5.3 Protection Is More Than Authentication

Throughout the analysis, it became apparent that most participants who did not use authentication to protect their phone did not consider themselves to be unprotected. We were able to identify a fair number of approaches that participants applied to protect their devices in the online study. These users felt secure despite the absence of authentication. For instance, participants reported to never leave their devices unattended when in public settings and to keep them close at all times (e.g. in their

[3] The finder is not able to access the address book to find the owner.

pockets or bags). This is also mirrored in the low number of high impact unwanted access possibilities during the field study.

This is even more interesting when analyzing the risks related to smartphone use. Only 26 % of the perceived worst-case risks in the study could actually be avoided by authentication. These included risks like theft or loss of the device itself. In many cases, participants rated the monetary value of their devices higher than the possibility of losing their data or someone gaining access to the data. Similarly, absence of threat was a very frequently mentioned reason for not having a lock screen in the online and the field study. This finding again provides evidence for considering user behavior as well as the ecosystem in which a measure will be applied when evaluating whether or not a measure has a chance to find adoption.

5.5.4 Sensitive Data is Seldom Accessed

As mentioned before, when filling out the questionnaire, participants were asked whether the accessed data is sensitive for them. This was the case in 25.3 % of all sampled unlocks. This means that nearly 75 % of interactions with the smartphones were with non-sensitive data. Taking into account the overhead created by the authentication process, there is high potential for lowering the burden for the users. That is, the results indicate that binary authentication as we are using it today (i.e. all or nothing access to a device) should be seriously re-assessed. For instance, instead of protecting the mobile operating system in its entirety, protection might be used on a data level. We can see a current trend in the mobile phone industry, granting access to non-sensitive functionality like flashlight and camera (not photos) without the need for protection. Our results suggest that this does not go far enough and more aspects of the phone could be used without the need for authentication. Hayashi et al. [80, 82] already proposed potential solutions for this problem, which can be able to reduce the usability impact of protection measures and hence lower the burden on users.

5.5.5 Shoulder Surfing Risks Perception

The results of the field study indicate that the perceived shoulder surfing risks are rather low. Our participants believed shoulder surfing would have been possible in 17 % of reported cases. However, it was considered a high risk in only 11 out of 3410 occurrences. Additionally, participants protected themselves against such attacks using physical measures only in 2.8 % of sampled situations. Overall, we can state that the participants were aware of possibly risky situations but that this did not influence their general opinion about protecting against this threat. While shoulder surfing can take place in any environment, unknown attackers are mostly present in public environments, which were however frequented least by our participants.

Shoulder surfing in private environments was mostly considered possible by people known to the user. This was, however, often not considered a threat or those people knew the lock codes anyway. Yet, this does not mean that shoulder surfing is not a

risk worth addressing by improved technology. Just because users do not perceive a threat as serious does not mean that it is not. It does however mean that the additional effort a user is willing to invest to protect from it needs to be carefully assessed. Based on our results we also recommend that the shoulder surfing attack risk can be minimized by reducing the number of "unnecessary" code entries. Since shoulder surfing resistant authentication mechanisms often incur reduced performance, the user should be able to decide in which situation protection is actually necessary. This was, for example, proposed in the XSide concept of De Luca et al. [41]. Also, testing shoulder surfing resistant mechanisms on an audience that may not see an immediate threat from such attacks should be carefully reconsidered.

5.5.6 Combining Activity Logging and Situation Sampling

Finally, the combined approach of using activity logging and experience sampling was well received by our participants. They reported no undue burden, felt that the collected data represented their behavior well, and even indicate to have benefited from the awareness gained through their participation. Sampling situations allowed us to gather diverse insights into participants' smartphone use while the logged data provided us with a solid foundation of participant behavior.

5.6 Limitations

The online study as well as the field study both have limitations. The online survey relied on self-reporting and can hence only shed limited light on real behavior. We therefore focused this investigation on respondents' perceptions, attitudes and common practices. The field study also logged behavioral data, but uses a different sample of participants as well as a limited set of sampled situations. While a considerable number of situations was sampled across 27 days, participants also indicated that some rare occasions and situations that did not last very long have been missed. Furthermore, extreme situations caused participants to dismiss the questionnaire, as they needed to access information quickly. Showing the questionnaires also heightened participants awareness of risk and their own behavior. This may have influenced participants' responses.

Similarly, we were only able to extract certain events from the Android OS. The reported times for the duration of the unlock therefore also include occasions where participants first read their notifications and only unlocked afterwards. The reported times should therefore be treated as upper limits. However, as this behavior is likely similar across the lock mechanisms, the respective values should still be comparable.

Finally, the field study also included self-reported and subjective views. Participants may have categorized similar situations as, for example, public or semi-public environments, depending on their perception. Also, the same data may be perceived as more or less sensitive by individual participants and attack opportunities may have been missed. However, we argue it is the participants' views that count more than

absolute numbers, as they are more likely to adopt improved security measures if they see a relevant threat by themselves. Participants may also have chosen that they do not know what kind of data they want to access when they were unsure about that data's sensitivity.

5.7 Summary

In this chapter, important insights into users' behavior and perceptions with respect to the use of a commonly used security measure were described. The study my colleagues and I conducted yielded deep insights and provided new ground truth for the future development of security measures, especially for mobile devices. The online survey gave a broad overview of participants' reasons for (not) using lock screens, how they protect their phones, and which critical incidents have previously happened to them. In addition, the longitudinal field study captured one month of unlocking activity and sampled 6582 situations in situ.

Main findings include that there is a massive number of unlocks that the participants themselves severely underestimated. Participants also showed very diverse reasons for locking or not locking their phone. We also demonstrated that users apply many physical measures to protect their phone, which often makes additional IT measures superfluous in the their opinion. The participants in our study indicated to access sensitive data in only one quarter of cases, which provides an opportunity to reduce the attack surface of shoulder surfing through context sensitive unlock mechanisms. The results again show the importance of considering users' self-efficacy when trying to find adoption for a new security measure. Also, as shoulder surfing is not a frequent threat and is rarely considered likely and severe, an absence of a perceived risk makes countermeasures less attractive for users. Finally, increased authentication time of a new authentication measure for mobile devices of only very few seconds can severely impact the perceived effort and therefore reduce users' readiness to adopt.

6 The Adoption Budget – On the Role of Risks and Consequences for Security Technology Adoption

The previous chapters shed light on factors influencing concrete security measure adoption decisions before or during their use. I was able to show that protection is only deemed necessary if a particular benefit is perceived by users. In this chapter, I investigate if a lack of risk and consequence awareness in general may be able to explain the lack in security measure adoption that is often observed. To this end, I extend the compliance budget theory to home users and provide evidence for the involvement of risk and consequence awareness and perception based on the results of a survey on users' risk perception during everyday Internet use.

6.1 Motivation

The theories and related work presented in Chapter 2 as well as the three studies I discussed in the previous chapters clearly show that users will only adopt a security measure if there is a clear benefit. Demonstrating such a benefit is particularly difficult for creators of IT security measures. As noted before, early research of Sasse et al. [140] established that, in contrast to other branches of usability research, usable security measures are particularly difficult to create, as security is rarely a primary goal or task for users. Additionally, not only do HCI researchers need to design security and privacy measures that enable people to remain safe while following their primary task, they also need to create a wish for adoption of these measures beforehand:

> "The challenge is not to enable the individual's mastery of an application so much as to convince the individual to avoid digital risks by adopting appropriate security tools and application settings, despite the financial and time costs of doing so." [14]

On top of the usability and environmental factors I have already investigated in the previous chapters, the problem with convincing users of security risks as well as the interplay with financial and time costs of security measures has been investigated by Beautement et al. [7]. According to this work, users' *compliance budgets* are limited and therefore users make a rational choice when they reject (new) security measures if they do not perceive enough benefits. The compliance budget was originally defined in a corporate context, where security policies are forced on users, who may

then choose not to comply. As already noted in Chapter 2, Herley [85] extended this reasoning to regular users and states that there is too little benefit while costs are too high when it comes to existing measures, such as SSL warnings or stronger passwords.

The reasoning behind the compliance budget can also be extended to the regular Internet user at home by asking the question why those users do not adopt security measures that can protect them from the many risks on the Internet. While there are no policies forced upon home users by a governing body, they still do mind their security to some extent and use some of the available measures. It appears that users do indeed have a variant of the compliance budget that lets them choose some measures but not others. In line with the previous term of Beautement et al., I propose the term adoption budget, as users do not have a set of organizational policies to comply with.

Consequently, the question arises why certain measures, such as choosing more difficult passwords for banking accounts, are currently within users' budgets. As there is no mandated policy to comply with, one can think of users spending their adoption budget on security measures as they seem fit. Assuming that users spend their budget on the most important and salient risks more easily, then there may be some risks users could be protected from but do not care enough to take the necessary precautions and some risks users would like to be protected from but aren't at the moment. Similarly, by looking at salient risks, I may find why some protection measures are currently not within the adoption budget: if a risk is considered unimportant or entirely unknown, users will not have any desire to protect themselves against that risk.

For security technology, previous research on technology acceptance models (e. g. [38, 83], cf. Section 2.3) also indicates that a subject needs to be aware of a threat – or risk – and then come to the conclusion that this threat needs to be dealt with. Only then is a security technology evaluated for its suitability, for example its usability and capability to protect against said threat. Corroborating these theoretical models, participants' comments in the study presented in Chapter 3 suggested that seeing personal consequences arising from a certain risk is an important factor when considering the adoption and use of a novel security technology.

Interestingly, to the best of the my knowledge, little is known about which IT security risks users actually feel exposed to and are aware of during their everyday dealings with today's Internet. Previous research on risk perception has often focused on describing specific threats before analyzing how users perceive those. Similarly, representative national surveys, such as the Oxford Internet Survey for the U.K. [50] or the Security Report for Germany [151], only ask participants about attitudes towards enumerated risks. While this gives an understanding of how users perceive a specific risk when prompted, these results can not give insight into the risks actually perceived in the users' everyday lives. As users are no security experts, they can and will only pay attention to risks they are aware of. Additionally, users may evaluate risks differently from security professionals, for example by considering other threats or reasoning differently about possible consequences. There has, however, not been

any recent work addressing this bottom-up view of risk awareness. It is conceivable that this lack of knowledge is detrimental to the foundation on which security and privacy researchers base their work. I argue that it is important to know which risks users are aware of and how they appraise them, since they will only take action based on their own, intrinsic appraisal.

Another factor that influences the perception of risk and therefore IT security technology is its perceived benefit. Beautement et al. [7] mention avoiding consequences and punishment as key benefits, diminishing the costs a mechanism creates within the compliance budget. Outside a corporate environment (in which Beautement et al.'s reasoning was situated), negative consequences are the only kind of punishment a users has to fear, as there is no regulatory body punishing misbehavior. Similarly, the technology acceptance model of Herath et al. [83], also features an assessment of effectiveness – i. e. is a technology able to prevent certain consequences from happening – in the appraisal of security technology. Hence, only the prevention of certain relevant consequences using a certain behavior or technology will constitute a benefit for users that can make them accept a certain cost in terms of effort. I therefore posit that without relevant consequences perceived by users, they are unlikely to adopt security measures or change behavior to guard against risks, even if the risks themselves are known.

The contents of this chapter contribute to providing a foundation on which to base future developments of security measures. The results can help researchers and developers target the risks actually perceived by users in given situations and also highlight how security measures need to address the most important risks perceived by users. By looking at the perceived consequences, I also analyze if there are risks users are aware of, but against which they do not protect themselves, because they do not perceive any relevant benefits. Overall, I argue that knowing which risks and consequences users want to be protected from is an important precursor for the adoption of measures or behavior change.

In the remainder of this chapter, I give an overview of users' risk awareness while using today's Internet, based on a survey of 210 participants from two different populations. The results suggest that the sporadic and slow adoption of new security measures, for instance replacing username and password, is not only due to usability problems but is also rooted in these measures not addressing salient risks. I also find evidence that there are some culture-specific risks, but many risks are also common to all participants: Malware, hackers and stealing account credentials were the most salient risks and financial losses accounted for the most frequently perceived consequences. I also compare the awareness of salient risks to agreement with a set of risks commonly warned against: the results suggest that users are aware of far fewer risks than may currently be believed. Overall, the lesson to be learned from this study is that there appear to be two avenues for improving end-user security in the future: One, accepting the limited set of salient risks users perceive (or wait for this state to change on its own) and create security mechanisms that address these risks while leaving users insecure against risks they do not care about. Two, supporting

the process of changing risk perception, for example using risk communication and education.

Disclaimer: The contents of this chapter were previously published and presented at the Computer Security Foundations Symposium (CSF) under the title "Who's Afraid of Which Bad Wolf? A Survey of IT Security Risk Awareness" together with Sascha Fahl and Matthew Smith [75]. All parts of this work were primarily completed by myself. The co-authors advised me on the study design and provided feedback on the manuscript. Sascha Fahl also helped during the coding process of the unstructured data.

6.2 Background and Related Work

In 2002, Friedman et al. [66] presented a short paper on a study of users' concerns about risks during the use of "the Web". The focus of their study was to analyze the effect of communities on risk awareness. They interviewed 72 people and found that participants "most often emphasized security, privacy, and threat[s] to computer systems" as potential risks. The more "high-tech" a community was, the more they were concerned about security and privacy. Additionally, users from a suburban community were more concerned about people and their experiences, while a rural community showed significantly less concerns overall. These results were obtained when the Web was relatively young and its use was not as common as it is today. Unfamiliarity and novelty has likely played a role in the participants' views more than 10 years ago and the authors argue that the investigated communities will likely progress towards the views of the high-tech community. Since this study was conducted, interaction with technology and the Internet has changed significantly and it is hence important to draw a current picture of which risks people see for themselves in this changed environment. I extend the study design of Friedman et al. in an effort to create a recent and more detailed understanding of people's general risk awareness when using the modern Internet with a focus on IT security risks.

As mentioned in the motivation section, there also are several representative, national surveys compiled on an annual basis about views on IT security and risks (e. g. [50, 151]). These ask their participants about their attitudes towards certain enumerated risks, for example having their computer infected with malware or being under surveillance by governments. This method works well to judge the relative relevance of different sources of risks, but does not address which risks may actually influence day to day behavior. However, models for technology adoption suggest that users will only consider measures against certain risks for their adoption budget if they are intrinsically aware of these risks and hence consider them worth their time.

Also, in the papers introduced in Section 2.4, risks were assumed to arise from specific threats, such as phishing, hackers, or malware, ignoring the fact that users may not be aware of such threats or believe that they do not apply to them. Threats

were always simply presented to the users in these studies. For users to adopt security measures protecting them from these threats, they will have to first discover a measure and then decide that it is worth the effort to actually use it. This appraisal, using the model of the adoption budget, will only have a positive outcome if a benefit, protecting oneself from a risk and its consequence, is perceived. However, in the study I presented in Chapter 3, I found that participants do often not differentiate threats or risk at all. Participants were confident that using only two different passwords across all their online accounts was safe and saw no problems or risks arising from that practice. Conversely, multiple participants also expressed that they treat the Internet as a generally insecure medium and that they therefore, for example, do not use online banking at all. Participants also expressed doubts that security technologies would actually protect them from the risks they perceive. Among other comments, one participant believed that password managers "*surely could be hacked by someone*". Another participant said: "*I don't believe that there will ever be perfect security on the Internet. Whether you use [an alternative mechanism] or continue using passwords [...] there are vulnerabilities everywhere*". In a study I conducted with Sascha Fahl [56] (which was also part of the same paper that featured the contents of Chapter 4), many subjects believed that there will be a way to circumvent any security system at some point in time and that virtually anything on the Internet was vulnerable to attack or "hacking". Similarly, Klasnja et al. [105] found that users "lack understanding of important privacy risks" when connecting to and using Wi-Fi networks. This corroborates the intuition that a specific risk as well as concrete benefits need to be perceived in order for users to consider countermeasures.

A non-tech oriented study by Hogarth et al. [88] investigated everyday risk perception. Participants recorded one risk and the most severe consequence involved in whatever they were doing when receiving a text message from the researchers three times per day on work days over the course of two weeks. They found that the most frequently reported risks were the most salient ones as opposed to the most severe ones. Consequently, they conclude that the risks users are aware of are only a subset of the risks actually faced. Many of the everyday risks commonly studied were also entirely absent from the risks reported by their participants.

To the best of my knowledge, there has not yet been a recent investigation of the risks and consequences perceived during everyday Internet use without being queried about risks arising due to specific threats. Additionally, this is the first study to analyze risk awareness with a focus on IT security in the modern Internet, providing ideas to facilitate the adoption of security measures.

6.3 Online Risks Survey

To investigate everyday risk and consequence awareness of users, I designed a questionnaire and ran an online survey. In contrast to Friedman et al. [66], I chose a survey as research method, because surveys can reach people in familiar settings.

While it is well known that using surveys to ask people about past behavior causes biases, my survey used scenarios to get people into a certain mindset before eliciting their attitudes within that mindset.

I was also concerned that inviting people to interviews may cause biases. For instance, participants who do not see many risks may feel obligated to name more risks in order not to seem careless. It also seemed likely that participants would not share risks they are afraid of because they might feel ashamed. It has been shown that responses are more truthful and open when given in private [141]. Yet, interviews can yield deeper insights into users' reasoning if care is taken to minimize bias and can therefore be complementary. I thus decided to use a survey method to gain a broad overview and conduct interviews in future work to explore individual aspects of the survey results in more depth.

I also chose a survey over more fine-grained methods, such as the experience sampling used by Hogarth et al. [88], as I posit that it is necessary to gain an initial understanding across a wider range of Internet users. Especially differences in culture and beliefs may influence risk awareness which I would not be able to capture using other methods. To investigate the adoption of security technologies on the Internet's scale, a wider view is important. I therefore ran the survey on two continents, using a local student population in Germany as well as workers from Amazon's Mechanical Turk. The questionnaire design and participant demographics are detailed in the following subsections.

6.3.1 Questionnaire

The questionnaire was structured to elicit a set of risks to which participants believe to be subject to during their daily Internet conduct. I presented five scenarios in which participants were asked to list which risks they are aware of. These scenarios comprised "using the Internet in general", "logging in to your social network account", "shopping online", "online banking", and "finding a shared ride using online services". The reasons for including each of the scenarios are as follows:

- *General Internet use:* This scenario was chosen to induce as little priming as possible to try and capture the base line of risks users are aware of. I hypothesized that users may apply this mindset when considering general protection measures without a specific application. The remaining scenarios include more concrete, yet common online use cases.
- *Online shopping and banking* were chosen since they both include obvious financial risks. I chose two financial scenarios to examine whether the type of institution influences the risks and consequences users state. While banking constitutes a more severe scenario, it could also be perceived to have less risks, as banks take more precautions to protect their customers.
- *Logging in to a social network site* was chosen because it is a very common activity and is often paid little attention to or even perceived as annoying. It suffers from the common problem of IT security mechanisms, since it is a

barrier keeping users from achieving their primary goal, which in this case is to take part in a social network site. This scenario was also chosen, as social networking accounts often hold more sensitive data and therefore potentially have a different protection value than for example credit card details in the shopping scenario.

- *Sharing a ride using online services* was chosen because it includes direct real-world implications, as the user will meet with another person in the physical world. I included this scenario to capture a potential relationship between real-world, physical risks and abstract, technological online risks.

Each scenario was introduced to participants with a short description. The description was brief and simple and aimed to let the participants imagine how they usually interact with such services. The text mentioned a well-known workflow for each scenario and reminded participants to imagine that they were completing this task in a familiar environment. This part of the description aimed to overcome the issue of trust, as they should trust their favorite shopping site, social network or bank similarly. For example, the description of the shopping scenario read:

> "Please imagine that you are using the Internet at home as usual. You are visiting your favorite online shopping site and would like to make a purchase. You enter your address and payment information on the site and complete the checkout process. Please answer the following questions in the context of purchasing merchandise in an online shop."

For each scenario, I first asked participants to state the most severe risk or danger they believe to be subject to within this scenario. To let participants state whatever comes to their mind first, the questionnaire provided text boxes for free-text answers. The text boxes were sized to accommodate approximately one sentence in one line so as not to overwhelm participants. I specifically chose not to define or otherwise explain the concept of risk, as I intended to capture how participants intuitively respond to the questions. I posit that if I asked participants to state a risk according to a certain definition, rationalization would overlay their initial responses as they would think too much about their answers in the context of a given definition. While I acknowledge that this lack of specificity may cause some risks which are technically similar to be described in different ways, this approach prevents that I miss differences in users' reasoning due to forcing them into a certain definition. Since these differences are one of the main focuses of the work presented in this chapter, I chose to accept this limitation as I mainly use the results discuss which differences in risk perception may be observable.

Also, one of the very fundamental techniques taught to security professionals is to evaluate and rank risks by combining the likelihood of a risk and the severity of consequences. Security professionals are trained in differentiating these terms and making decisions based on the technical understanding and hopefully well founded experience. If the general population does not make the same distinction in these terms, their basis for making decisions is different. I believe that examining the differences in awareness and perception is a vital foundation to understanding users

and their decisions better. As argued above, I would not have been able to examine this important difference if I had provided participants with expert definitions during the survey.

Next, participants were given the chance to enter three additional risks for each scenario, before being asked when they had last heard about each of these risks from common sources, including friends, family, and media. I then asked participants to rate the completeness of their set of risks and to give an estimate of relative risk arising to their wellbeing in general on a scale from 0 (no risk at all) to 100 (greatest possible risk).

Participants were then requested to state four potential consequences in order of severity for each scenario. As before, I chose open-ended questions with free-text answers so as not to influence their answers. As before, I also did not provide a definition or explanation of the term consequence, in order to preserve the participants' mental models as much as possible. I asked for consequences, as this will allow me to better assess to what extent users conform to the common model of security professionals and also to what extent they are actually ready to do something against a risk or at least to what extent they believe a risk applies to them. Based on the work presented in Chapter 3 as well as the compliance budget theory [7], I assume that if users see only very improbable or impersonal consequences, they are very unlikely to see a need for measures against the corresponding risks. Therefore, participants were also asked to judge the relative severity of the most severe consequence in comparison with those arising from other risks and dangers in their life. Additionally, each scenario concluded with a question about the perceived likelihood of the most severe consequence happening to the participant personally. An overview of the questions included in the questionnaire for each scenario can be found in Appendix D.1.

The questionnaire ended with a block of questions giving the participants a precompiled list of 22 common risks users are often warned against or which featured on popular websites about online risks, asking them which they know about and how relevant they consider those on a scale from "not relevant at all" (1) to "very relevant" (4). It included common risks like malware, spam, phishing or online shopping fraud, but also other dangers, such as psychological issues due to exposure to unsuitable content, cybermobbing or Internet addiction. In this part of the questionnaire I intended to compare the open-ended answers from the first part to the risk users would say are relevant or they know about if the risks are presented to them in a list. Users' views on risks are often collected in this fashion in representative surveys (see above). I believe that this introduces biases that makes results based on enumerated risks less suitable for analyzing why security measures are adopted or not.

The questionnaire also asked participants about their perception of risks and security on the Internet in general, how much which sources of risk information influence their perceptions, and if they had previously been subject to any of the consequences and risks they gave before. Lastly, demographics were collected.

6.3.2 Participants

As mentioned above, I recruited students from my university's study participation mailing list and submitted a task to Amazon's Mechanical Turk in July 2013. For the university students, the survey was administered in German. The students were offered to enter a raffle of 30 10 Euro Amazon vouchers. On MTurk, I invited only U.S.-based Master workers to work on the task, offering $ 3 for a 20-30 minute survey. Amazon screens Master workers for reliability and they receive a higher compensation per task. This should diminish the impact of workers trying to make as much money with as little effort as possible on the reliability of the results. I still checked all results for irregularities and obvious patterns in the answers and removed one participant who was answering randomly.

The choice of participant recruitment offers a look at two rather different populations and allows to look for major differences in risk awareness between these different countries, education and age groups. While the two samples do not represent the countries in general, they do offer a fairly broad view across a diverse set of people.

I received $N_1 = 111$ complete questionnaires from the university students and $N_2 = 99$ complete and valid questionnaires from MTurk. Students spent 24.7 minutes ($sd = 16.5$ min) and Turkers 22.7 minutes ($sd = 11.3$ min) on the questionnaire. Table 6.1 gives an overview of participant demographics. MTurk workers were older and comprised more females. Their IT experience was similar while previous experiences with online risks and dangers was reported to be higher by students.

	Students	**Turkers**	
N	111	99	
Age Range	18-42	19-66	years
Median Age	23	36	years
Gender	45.0 %	60.6 %	female
	55.0 %	39.4 %	male
Occupation		37.4 %	full-time employee
	100 %	6.1 %	student
		11.1 %	part-time worker
		20.2 %	self-employed
		11.1 %	homemaker
		9.1 %	unemployed
		5.1 %	retiree
IT Experience	21.6 %	18.2 %	is currently or has been working in or studying IT
Risks	59.0 %	35.4 %	previous incidents
	6.7 %	4.0 %	N/A

Table 6.1: Participant demographics for both survey deployments.

Differences between Students and Turkers

As I aim to investigate risk awareness for an as broad as possible population and to maintain a certain level of clarity in the results, I combine the two datasets for the analysis I present below. Whenever there were significant differences in the results for a certain aspect, the respective results subsection includes a description of how the two populations differed. Otherwise, the conclusions drawn from the data apply to both populations equally. The differences I found will also be discussed in the Discussion section of this chapter.

6.3.3 Coding

To analyze participants' responses to the open-ended questions on risks, I used an inductive coding procedure. I chose this method to be able to flexibly represent the responses, as there is, to the best of my knowledge, no previous research on fine-grained coding and categorizing for user risk awareness concerning the Internet.

Coding began with the list of 22 common risks that was also included in the later part of the questionnaire. One coder went through all 4,200 responses, adding codes whenever an answer did not match an existing code. Codes were also hierarchically refined if a response fitted an existing code but addressed a more specific aspect. Each code could only be assigned once per user and scenario and was otherwise marked as duplicate. Coders also filtered for responses that were not descriptions of a risk but of a consequence or something else. After this first coding session, codes were refined in a discussion among me and my co-authors and a second coder went through the responses again using the refined coding scheme. The same process was applied to the open-ended responses on possible consequences.

Overall, 74 risk codes and 38 consequence codes were created. Codes that address a common concept were grouped into a hierarchy. A table of all risk codes and their hierarchy can be found in Appendix D.2 together with the counts for each code. I did not consolidate the codes further for the purpose of this work, as I aim to explore risk awareness and its implications for the human factors of the adoption of security measures on a broad scale.

6.3.4 Results

Altogether, 210 participants had the chance to specify 20 risks each. Of these potential 4,200 risks, a total of 1,795 valid responses across the five scenarios were given (median of seven unique risks per participant). The remaining responses were either empty or filtered for several reasons (see below). Figure 6.1 provides a graphical overview of all mentioned risks, grouped into the categories that emerged during coding. In general, concerns about privacy, account abuse, malware and hackers, as well as financial risks and fraud were most commonly mentioned. There is also a fair amount of miscellaneous risks that were mostly mentioned in the ride-share scenario, as "unreliability of people" was a frequently stated risk in this case. Since

Figure 6.1 also suggests that risk awareness depends essentially on the presented scenario, I look at the scenarios individually below.

Differences Between Populations I performed a Fisher's exact test over the data source × risk-code contingency table. As the table was too large to compute all permutations and had more than 70 % of expected counts at less than five due to the sparse nature of many risks participants specified, I used Monte-Carlo simulations to obtain an approximate p-value. The test indicated a highly significant difference between the two data sources ($p < .0001$, 100,000 replicates). Significant standardized residuals in the contingency table revealed differences within the following codes (all these cells had expected counts larger than 5): Turkers more frequently gave identity theft, abuse of bank details, and theft of physical items as possible risks. The German students more often named fraud, hidden cost, and abuse of personal data (cf. the counts in Appendix D.2).

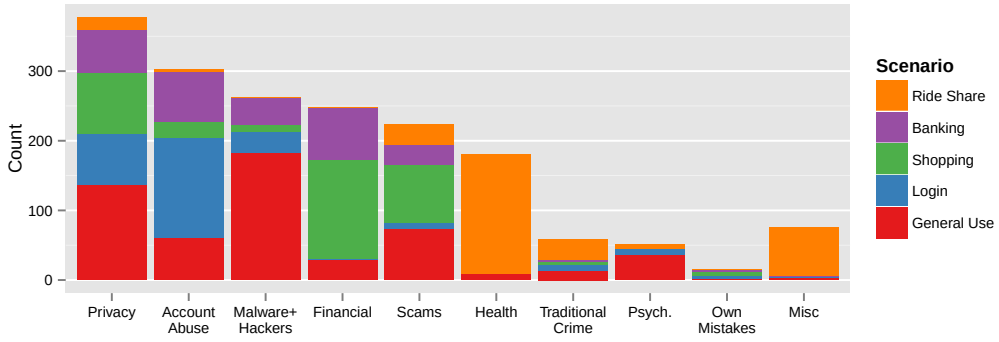


Figure 6.1: An overview of all risks mentioned by respondents grouped by categories and color-coded by scenario (cf. Section 6.3.1). Counts for all risk codes can be found alongside the codeplan in the Appendix. The counts of mentioned risks differed significantly from the expected counts for the categories privacy (students 220/turkers 157), scams (106/118), health (78/102), and traditional crime (12/47) due to the differences in mentioned risks detailed in Section 6.3.4.

Scenarios

Table 6.2 provides an overview of top risks by presented scenario, including these risks where the populations differed. The following results nevertheless largely apply to both populations. In the general use scenario, I find that users are most concerned about malware, identity theft, stolen account credentials and hackers. The lesser concerns mostly comprise privacy risks. The great fear of malware is interesting, as a plentitude of software exists that can help users to protect them from this threat, yet users apparently still feel that they can be subject to adverse effects from it. I suspect that participants may nevertheless be afraid to unknowingly contract malware due to a lack of understanding for technical complexities. The fear of hackers possibly

arises from a similar source, as participants may believe that a competent person can break into almost any IT system.

Risk	Count	S/T	MS Count
General Use	**548**		**205**
Malware	121		42
Identity theft	55	8/47	28
Stealing account credentials (specific)	46		5
Targeted attacks by third parties (“Hackers”)	42		18
Stealing private information	41		22
Loss of privacy in general	26		13
Surveillance	23		12
Abuse of credit card/banking details	21		10
Abuse of personal data	15		7
Stalking	11		1
Shopping	**358**		**200**
Abuse of credit card/banking details	125		108
Fraud	35		10
Stealing private information	34		13
Passing private information on to third parties	26		13
Offering non-existent merchandise or services	21		7
Identity theft	16	0/16	6
Stealing account credentials (unspecific)	11		6
Banking	**281**		**184**
Abuse of credit card/banking details	58		49
Stealing account credentials (unspecific)	33		28
Stealing account credentials (specific)	29		24
Stealing private information	24		17
Identity theft	23	1/22	7
Targeted attacks by third parties (“Hackers”)	22		20
Surveillance	12		7
Malware	12		3
Login	**278**		**184**
Stealing account credentials (specific)	67		60
Stealing account credentials (unspecific)	64		59
Stealing private information	19		10
Targeted attacks by third parties (“Hackers”)	16		12
Surveillance	12		5
Ride Share	**330**		**195**
Risk to health and wellbeing	157		131
Unreliability of other people	69		19
Theft of physical things	19	0/19	0
Fraud	17		15

Table 6.2: Risks mentioned by more than 5 % of respondents grouped by scenario. The MS Count column lists how frequently this risk was stated as the most severe risk (being elicited first in the questionnaire). The S/T column lists counts for the student (S) and MTurk (T) populations separately if the counts differed significantly between them.

Looking at the top risks mentioned in the four specific scenarios, I find that the more technical risks of the general scenario are superseded by risks that mirror the described scenario. When considering shopping online or online banking, risks pertaining to the respective task, such as abusing account details or fraudulent merchants, were considered to be most important. This confirms the common view that security is a secondary issue, even with respect to risk awareness. Notably, stealing account credentials was more relevant in the Banking scenario, even though banks usually take greater care to protect their customers.

In the scenario that specified an IT-security relevant activity (logging in to a social network), I find that IT security risks were considered most important again. In consequence, this could indicate that users only see a necessity for improved security mechanisms (that target IT security risks) when these risks are obvious in the corresponding situation and are not overlaid by other, more important, non-IT-security concerns.

There also are two risks that occur in most of the specific scenarios: identity theft and stealing private information appear to be concerns that are cross-cutting for most Internet usage scenarios. While private information often inherently needs to be entrusted to online services if they are to provide a useful service, identity theft could be more difficult if authentication of individuals used appropriate protection measures. This is also a risk that was particularly pronounced for participants from the U.S., as social security numbers are often used for authentication in important and official workflows in this country.

I furthermore compared the provided risks across all responses with the first response participants gave in each scenario. In the questionnaire, a scenario would be introduced and the participants were then asked to first state the greatest risk they thought may arise from this scenario. I found that while the top two or three risks provided based on the greatest risks and on all given risks respectively did not differ, there were some risks that seemed to only occur to participants on second thought (cf. third column in Table 6.2). These included stealing of account credentials using a specific means in the general use scenario, malware in the login scenario, fraud in the shopping scenario, identity theft in online banking, and theft of physical things in the ride share scenario. Hence, if users assess a security measure in a short amount of time, some risks may not be considered. Similarly, future studies that wish to elicit a comprehensive overview of specific risks should therefore plan to include more than one opportunity to specify a risk.

Filtered Responses

As noted above, participants had a total of 4,200 slots into which they could enter risks. 2,021 slots were left empty. Non-empty slots were filtered further: duplicates, where the same participant stated the same risk twice in one scenario, were found in 201 cases. In thirteen cases, participants stated that a scenario did not apply to them and in 29 cases they stated that they did not see any risk in this scenario. Finally, there were 21 cases where answers were considered off topic by the coders

Answer	General Use			Login		
	S	T	Overall	S	T	Overall
all relevant	7.3%	52.0%	**28.4%***	22.9%	56.1%	**38.9%***
most important	64.5%	27.6%	**47.1%***	32.4%	11.2%	**22.2%***
all known	19.1%	17.3%	**18.3%**	35.2%	29.6%	**32.5%**
feel safe	3.6%	3.1%	**3.4%**	3.8%	2.0%	**3.0%**
space full	3.6%	0.0%	**1.9%**	1.0%	1.0%	**1.0%**
no answer	1.8%	0.0%	**1.0%**	4.8%	0.0%	**2.5%**

Answer	Shopping			Banking		
	S	T	Overall	S	T	Overall
all relevant	20.0%	59.6%	**38.8%***	18.2%	54.6%	**35.4%***
most important	50.9%	13.1%	**33.0%***	45.5%	18.2%	**32.5%***
all known	24.6%	24.2%	**24.4%**	28.2%	23.2%	**25.8%**
feel safe	1.0%	1.0%	**1.0%**	2.7%	4.0%	**3.4%**
space full	0.0%	2.0%	**1.0%**	1.0%	0.0%	**0.5%**
no answer	3.6%	0.0%	**1.9%**	4.5%	0.0%	**2.4%**

Answer	Ride Share		
	S	T	Overall
all relevant	13.0%	51.0%	**31.1%***
most important	44.4%	16.3%	**31.1%***
all known	26.9%	25.5%	**26.2%**
feel safe	0.0%	3.0%	**1.5%**
space full	3.7%	3.0%	**3.4%**
no answer	12.0%	1.0%	**6.8%***

Table 6.3: Participants' responses when asked how complete they estimate the set of risks they provided to be. The table shows proportions of answers separated by scenario and population (**S**tudents, $N = 111$, and M**T**urk, $N = 99$). An asterisk denotes a significant difference between the two populations in this scenario using a Fisher's exact test and standardized residuals greater than 1.96.

and six cases where a participant was unsure about potential risks for a scenario. Interestingly, in additional 114 cases (74 unique users), I found that users specified things that did not refer to an actual risk. While this can be an effect of fatigue, it may to a certain extent also be indicative of users not thinking about risks like experts. I did not find significant differences in the counts for filtered responses between the student and MTurk population. All in all, I gathered 1,795 valid responses. This suggests that users are only aware of a limited set of risks: only three participants exhausted all 20 fields to enter risks.

Completeness

When asked to what extent participants believe their answers to be complete for each scenario, they showed high confidence in their answers: In 34.5 % of all cases,

the participants stated to be sure to have entered all risks that are relevant for them. In an additional 25.4 % of all instances, participants indicated that they did not know about any additional risks. In 2.9 % of all cases, participants preferred not to answer this question, in 1.5 % of all cases no additional risks were entered because the questionnaire did not provide more space and in 2.4 % of the cases no additional risks were entered because participants stated to feel safe on the Internet. In 33.2 % of all instances, participants admitted that they were aware of additional risks, but that they entered the risks most important to them. Table 6.3 provides an overview of the answer proportions by scenarios and shows differences between the populations.

It is evident that the student population more frequently admitted that they only stated the most important risks while there are more. This in combination with the fact that almost no participant exhausted all possibilities to enter risks into the survey shows a limitation in awareness combined with a view that there are additional, but currently unknown risks that is more pronounced in the student sample. The majority of participants from MTurk stated that they have entered all risks that are relevant to them.

With respect to differences between the scenarios, participants more frequently stated that they provided all the risks they know about or consider relevant (71.4 % vs. 56.6 % overall, Fisher's exact test, $p = .001$) in the more technical login scenario. This indicates that the more abstract nature of this scenario caused participants to be less confident about the completeness of the risks they are aware of.

Prompted vs. Unprompted Risks

Comparing the risks participants entered in the first, open-ended part of the questionnaire to the risks people stated to know of in the last part confirmed that biases and priming severely impact risk awareness results. While 74.7 % of participants stated to know 15 or more of the 22 listed risks, only 22.5 % gave six, seven or eight of the risks they indicated to know about in their free-text responses. The remaining participants had less matches and none but one participant had actually previously mentioned all risks that he or she selected from the list.

Among the risks selected from the list, abuse of login credentials was the most well known risk with 96.2 % selections and psychological issues due to unsuitable content was the least well known (49.5 %). The latter risk was also never stated by any participant without prompting. Note that this and several other risks were not mentioned at all in free-text responses, while all risks on the list were indicated to be known by at least about half of the participants. This highlights the importance of measuring risk awareness without prompting users about specific risks.

Figure 6.2 gives an overview of the risks participants were able to select from the list. I also asked participants to judge the relevancy of each risk to themselves personally as "not relevant at all", "somewhat relevant", "relevant" or "very relevant". The labels printed in bold in Figure 6.2 show the risks that were considered relevant

Figure 6.2: Percentage of participants who stated a risk without being prompted that they later also selected from a pre-compiled list of risks. Labels printed in bold show the risks that were considered relevant or very relevant by more than two thirds of participants.

or very relevant by more than two thirds of the users. The bars in the figure also show how many participants provided these risks in the unprompted part of the questionnaire and then also selected them from the list.

In this analysis, the two populations only differed mildly: seven of the items presented in Figure 6.2 switch places with their neighbors when looking at students only. A notable exception is identity theft, which was considerably less mentioned by students as already discussed in Section 6.3.4 above.

Last Information on Risk

When asked when they had last heard about the risks they provided, participants mostly relied on information that was several weeks or older. As Figure 6.3 suggests, risks for the general use scenario appear to be based on recent information, while participants indicated to have less recent information for the specific scenarios. Considering the results described above, depending on the scenario, participants can more easily rely on available information from friends, family and media. The more specific the scenario, the less information can be obtained from these sources. Furthermore, the additional risks (columns 2-4 in each scenario in Figure 6.3) appear to be supported by less recent information in several cases, suggesting that risks which

participants had recently heard of came to mind first. This suggests that the availability heuristic of Tversky and Kahneman [153] also plays a role in the appraisal of IT security risks. Furthermore, participants indicated to more frequently never having heard about risks for the banking and ride share scenarios. There were no significant differences between the student and MTurk population in this analysis.

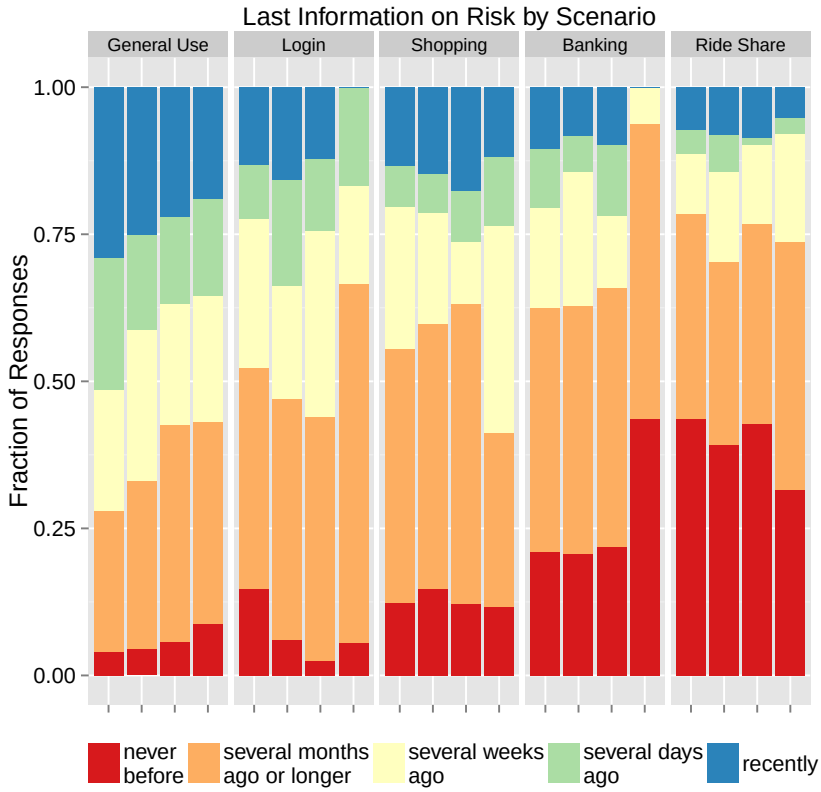


Figure 6.3: Overview of participants' responses to when they had last heard about the risks they provided, grouped by scenario. The leftmost bar in each group corresponds to the most severe risk stated by the users and the remaining three bars to the three additional risks.

Scenario Rating

Participants were asked to rate the relative overall risk each scenario poses to their wellbeing in general on a numeric scale from 0 (no risk at all) to 100 (greatest possible risk). Overall, all five scenarios were not perceived as posing a very high risk: The mean rating was between 31.8 and 35.5 (sd between 24.1 and 29.5) for all scenarios, except login, where the mean rating was only 25.1 ($sd = 23.8$). I conducted a 2×2×5 (data source × gender × scenario) mixed ANOVA on the participants' ratings and found a significant main effect for gender ($F(1, 206) = 6.87$, $p < .001$) and scenario ($F(4, 824) = 3.12$, $p = .023$, Greenhouse-Geiser corrected), as well as a significant interaction between scenario and source ($F(4, 824) = 3.30$, $p = .018$, Greenhouse-

Geiser corrected). The main effect of data source ($p = .33$) as well as the remaining interaction effects ($p > .26$) were not significant.

Looking at the effects, female participants rated the overall relative risk arising from the scenarios with 37.3 out of 100 while male participants only gave an average rating of 26.8. Furthermore, Holm-corrected pairwise comparisons revealed that the lower ratings for the login scenario were significant in all cases ($p < .05$) except when compared to ride-sharing, where ratings varied too widely. For the interaction effect, Holm-corrected pairwise comparisons showed that the login and general use scenarios were rated significantly higher by students ($p = .033$ and $p = .018$ respectively). Even though there were some significant differences between the ratings, the large standard deviations underline that the perception of risk severity varies strongly between individuals. Also, the relatively low ratings across all scenarios suggest that using the Internet is not a great source of risks for participants.

Risks and Security Measures

To see to what extent participants could have protected themselves against the risks they currently are aware of, risks were categorized into two classes: Those risks participants gave that can be directly addressed by a particular security measure, for instance malware infection, and those risks that cannot, for instance meeting with a serial killer. I then counted the number of times these risks were stated by participants in the general Internet use scenario. This scenario was chosen since I assume that a user would think about whether or not to adopt another security technology, such as a password manager, within this mindset, as many security measures are considered for all use cases a user may have online. If I then found risks that are currently addressable by security measures in the responses, this can indicate that the user does not feel adequately protected by the currently applied measures and thus possibly desires an improved solution. Also, I argue that the risks a security measure addresses need to be salient enough in their intended audience so that potential users feel a need for improved security through additional measures. Hence, the risks users were aware of and from which they could be protected from are in general good candidates for future developments of improved security measures.

In my sample, I found that 24.6 % of the risks mentioned by participants could be addressed by malware and spam protection, 16.4 % by privacy technology or legislation, 9.5 % by authentication technology, and 8.6 % by end-to-end encryption or access control. The latter proportion was significantly different between students and MTurk participants with 13.1 % of risks mentioned by students addressable by such encryption and access control measures and only 4.0 % of risks mentioned in the MTurk population.

Another possible explanation for this pattern in the data is the influence of advertising and media. The most frequently mentioned risks relating to malware protection are addressed by a flourishing market of protection products that many people use. Similarly, privacy measures have found a large audience in the press and politics, even before the recent reporting on the extent of the NSA's surveillance. Risks com-

monly addressed by the remaining security technologies – encryption, access control and authentication – correspondingly, were not as pronounced in the sample, possibly because they are subtle and not advertised by companies trying to sell a product. This may be an instance of how advertising and media coverage "educate" users to heighten awareness of malware and privacy violations as risks.

Overall, security-measure-related risks comprised 59.1 % of all risks mentioned by participants in the general use scenario (69.0 % in students, 49.3 % in Turkers). This also means that the considerable portion of 40.9 % of risks participants are aware of (31.0 % in students, 50.7 % in Turkers) can currently not be addressed by available security measures or privacy regulations and may hence induce a feeling of helplessness in users.

Consequences

Participants mentioned a total of 1,277 consequences across all scenarios. The consequences were elicited in order of severity. Table 6.4 shows an overview of the most frequently mentioned consequences. I find that participants mostly stated financial losses as potential consequences. The third most-frequently mentioned consequence "damage to health" mostly arises from the ride share scenario, where physically meeting with an unknown person was considered very risky by the majority of participants. Inconveniences, annoyances, and loss of time were also frequently mentioned among more serious consequences, including being victim of a crime or losing one's privacy.

Consequence	**Count**	**S/T**	**Unique Count**
Financial loss	283 (22.2 %)		145 (69.0 %)
Large financial loss	142 (11.1 %)		92 (43.8 %)
Damage to health	117 (9.2 %)		112 (53.3 %)
Inconvenience	81 (6.3 %)		65 (31.0 %)
Identity abuse	68 (5.2 %)		56 (26.7 %)
Victim of a crime	66 (5.2 %)	3.0 %/7.3 %	61 (29.0 %)
Loss of privacy	50 (3.9 %)		44 (21.0 %)
Annoyance (e. g. legal)	45 (3.5 %)		33 (15.7 %)
Loss of time	45 (3.5 %)		50 (19.0 %)
Loss of data	41(3.2 %)		50 (19.0 %)
Invalid: Not a consequence	576	212/364	
Invalid: Duplicate	130	42/88	

Table 6.4: Top 10 consequences mentioned by participants across all scenarios. The S/T column shows where the proportions differed between the student and MTurk population based on significant standardized residuals in the corresponding contingency table. The column Unique Count shows counts and percentages of unique participants stating this consequence.

Again comparing the two populations, damage to health, identity abuse and becoming victim of a crime were more frequently mentioned by participants from MTurk. The risk of identity theft was already more pronounced in turkers as described in

Section 6.3.4. The counts in Table 6.4 additionally provide some evidence that physical consequences are more pronounced in the MTurk population.

Table 6.5 shows the perceived severity and likelihood of the consequences, based on the instances where the consequence was listed as most severe. Interestingly, losing data was also mentioned frequently, but was not considered to be very severe, especially by MTurk participants. Generally, severe consequences were most frequently mentioned but also perceived as being rather unlikely. Losing privacy was the most likely consequence, according to the participants.

Consequence	**MS Count**	**S/T**	**Likelihood**	**Severity**
Financial loss	194 (30.7 %)		3.78	8.31 (8.1/8.7)
Large financial loss	122 (19.0 %)		3.79 (3.1/4.2)	9.01
Damage to health	112 (17.5 %)	12.8 %/23.0 %	3.37	9.03
Inconvenience	17 (2.7 %)		3.25	6.50
Identity abuse	31 (4.8 %)	2.3 %/7.8 %	3.42	7.71
Victim of a crime	22 (3.4 %)		3.81	8.52
Loss of privacy	12 (1.9 %)		6.00	7.33
Annoyance (e. g. legal)	6 (0.9 %)		4.33	8.33
Loss of time	6 (0.9 %)		4.00	6.83
Loss of data	5 (0.8 %)		4.20	3.80 (6.5/2.0)
Invalid: Not a cons.	242	83/159		

Table 6.5: Top 10 consequences mentioned by participants across all scenarios and their incidence as the most severe (MS) consequence and the corresponding percentage of valid consequences. The S/T column shows where the proportions differed between the student and MTurk population based on significant standardized residuals in the corresponding contingency table. The remaining columns show the average likelihood and the average severity assigned to this consequence. For both columns, participants gave rating on a 10-item numeric scale from "not severe/likely at all" (1) to "very severe/likely" (10). Rows with additional values in braces indicate values that differed significantly between the two populations (Student's t-test, $p < .05$).

Perception of Security and Information Sources

I asked participants for their agreement with three statements about their security when using the Internet on a scale from "do not agree at all" (1) to "completely agree" (7). 38.1 % of participants indicated that they often worry about risks and dangers in their day to day life at least somewhat (i. e., agreement > 4), 70.7 % of participants stated they feel at least somewhat safe when using the Internet and 31.1 % of participants at least somewhat agreed that they only have little influence on their security on the Internet. The agreement did not differ significantly between students and Turkers (Pearson's χ^2 tests, $p > .08$).

Considering sources of information for risks, participants were asked to indicate their perceived influence on a scale from "no influence at all" (1) to "great influence" (7). Media coverage was perceived to have more than medium influence (rating > 4) by 61.4 % of participants, 78.6 % stated this for stories told by friends and family, 72.6 % for the influence of information actively sought by the participants themselves

and 66.8 % for the influence of own negative experiences. This also supports the hypothesis that advertising may have influenced risk perception to some extent. Yet, social sources and own information gathering as well as experiences also appear to have considerable influence. It is noteworthy that the students reported significantly more high-influence ratings in all four cases (Fisher's exact test, $p < .028$).

6.4 Discussion

The results I obtained using this bottom-up survey method show very interesting aspects of risk and consequences awareness in participants, which are of interest for future attempts to communicate the benefits of security measures to users. On a very basic level, the results demonstrate that users differentiate between scenarios when assessing risk for things they do on the Internet. Many of them routinely consider multiple information sources, feel generally at least somewhat safe and believe that they can influence their security on the Internet. However, in contrast to the risks commonly addressed by security mechanisms, the risks participants are aware of are not very technical but of a more general nature, which needs to be taken into account when security experts try to address them. When other sources of risk in a scenario are more important than IT security (e.g., "Will I be physically harmed doing this?" or "Will the sweater look good on me?"), people might not think about other, IT-security-related risks. This does not necessarily mean that they are not aware of those risks at all, they are just not aware of them right then.

For example, only a single participant explicitly stated that he was concerned about man-in-the-middle attacks and this participant also self-rated himself as an IT expert. This is a particularly pertinent fact, since it is precisely this risk that the common SSL warning messages attempt to address. The results also suggest that a major factor is users seeing risks arising from the impersonality of the Internet: being a victim of fraud by unknown merchants or unreliable people, having a "hacker" attack one's accounts or data, and contracting malware from unknown sources or unknowingly becoming part of a botnet were frequently stated risks. Unknown attackers feature in all those risks and may therefore be believed to be hard to defend against.

Similarly, technical complexity appears to be a cause for concern in users: accounts being hijacked, credentials being "hacked" or stolen, and losing privacy were also commonly mentioned. What it really means to have an account hijacked or how credentials can or cannot get hacked is likely unclear to participants, as their often abstract and unspecific responses suggest. Also, how and to which extent a loss of privacy may occur in the scenarios was often not specified. Related work, for example the study of Rick Wash on users' mental models of threats [156], found that users often have incomplete knowledge of threats and underestimate the danger for themselves.

Finally, it is important to remember that participants were only able to state a median of seven risks of 20 possible risks. On top of that, the risk rating for all five

scenarios only yielded ratings of 25 to 35 out of 100, suggesting that risks arising from using the Internet are not a great concern. Thus, in terms of a security measure compliance or adoption budget, this may mean that a new measure either needs to address an already salient risk or find a way to raise awareness for its benefits while users' concern is low to begin with.

6.4.1 Asking Users About Risks

I was able to show that the set of risks users are aware of and that they can readily consider for their decisions is fundamentally different from the risks they indicate to know about given a pre-compiled list. Therefore, simply asking whether or not a user is afraid of a certain risk, scenario, or threat can generate misleading results. This is a very common practice in many studies and the results should be interpreted with care. Looking at the data collected in the study given the precompiled set of risks, many participants would have agreed that phishing, leaving a trail of data, and spam are relevant or very relevant risks. However, few participants actually mentioned these risks in the unprompted part of the survey. Previous work has shown that users don't readily engage with information in security decision situations and hence would also not be convinced that a particular risk is important in terms of their adoption budget. I therefore postulate that the set of salient and important risks a user is aware of mainly informs decisions in such situations.

6.4.2 Risks and Consequences

Considering the consequences participants see for the risks they specified, participants' reasoning appears to differ from how an expert would evaluate risks and their consequences. Participants often articulated something that is not a true consequence but a risk or simply a state of the world. For example, frequently, the consequence of an account being hacked was "my account is hacked". Similarly, "my credit card number is given to another person" was given as a consequence of the risk of disclosing one's credit card details. Thus, I argue that many participants do not evaluate which risks actually have tangible consequences for them and therefore underestimate the impact of a risk for themselves. If they do, the consequences appear to mainly relate to losing money, damage to their health and inconvenience.

At the same time, the low likelihood ratings indicate that they do not believe this will happen to them personally any time soon. A notable exception in the data is privacy, where the likelihood of having one's privacy compromised was considered fairly high. A potential moderator for these results is users' perceived self-efficacy, meaning how well users think they can protect themselves from a risk and its consequences. Even though I did not explicitly collect information about participants' perceived self-efficacy, it is conceivable that self-efficacy is low for protecting against a loss of privacy and therefore the likelihood of this happening to oneself considered high. Beyond an impact on likelihood ratings, perceived self-efficacy may have caused participants to not state some risks at all, as they feel that there is no threat arising

from risks they can cope with by themselves. Future work should explicitly look at self-efficacy as a moderator for risk awareness and hence the adoption of security measures.

During the process of coding the responses, coders noted that many, especially non-financial consequences were phrased in an impersonal way, for example "data stolen", "loss of privacy", "losing friends". While this might very well be a grammatical oversight or abbreviation, the frequency with which a mixing of personal with impersonal statements by the same participant was observed, provides some evidence that this may have an influence on risk and consequence perception. I thus posit that impersonal consequences cause some risks to be ignored, as the consequences are not perceived to apply to oneself personally and therefore remain abstract. Alternatively, in terms of the adoption budget, the cost of protecting against these risks may not be worth the potential benefits, as the consequences are too abstract to be of sufficient value.

The results show some important areas where an experts' view and participants' view of risks and consequences differ considerably. Especially the inability to see personal or any relevant consequences may influence a user's view on the necessity of adopting security technology or behaviors. I argue this represents valuable information, which needs to be taken into account when designing new IT security solutions and risk communication methods for end-users.

6.4.3 Differences in Risk Awareness

The choice of scenario influenced the set of perceived risks in participants. In the banking and shopping scenarios, financial risks were a lot more important, while risks to health and wellbeing overlaid many other risks in the ride share scenario. The study hence confirms that IT security often plays a secondary role in risk awareness when real-world risks are involved. Also, this means that the usage context of a security measure can influence its appraisal, as some risks become less salient in the light of other problems with regard to a certain task.

The results also indicate differences in risk awareness between the two populations that were used in this investigation. Workers from Amazon's MTurk, who were all based in the U.S. according to MTurk's filter, reported to be older than the student sample and included mostly non-students. These participants had a greater fear of identity theft, possibly because of the reliance on social security numbers and credit scores for many important financial aspects, which play a much smaller role in Germany. Similarly, fraud and scams in online shopping as well as hidden costs in services is a common concern for the German student participants. Furthermore, consequences in the real world were also more pronounced in MTurk participants, as they more frequently reported damage to their health and becoming victim of a crime as consequences arising from the risks they provided. I hence posit that taking these kinds of differences in risk awareness into account when designing security measures and developing strategies to deploy or advertise security and privacy measures can have beneficial effects.

Despite the Internet being used by a very diverse population, it seems that many security and privacy mechanisms are currently deployed on a global scale irrespective of culture or background. Future work of the usable security community could examine the benefits of tailoring design, presentation and deployment of security and privacy mechanisms to different cultures. Additionally, I found that the relative severity of risks arising from the very general and common scenarios varied widely between participants. Female participants also found the scenarios to be more severe than their male counterparts, which may also be a good target for tailored solutions.

6.4.4 Awareness of Own Negligence and Mistakes

Another notable result of the survey is the almost total lack of awareness for risks arising due to own mistakes or negligence. Only in eight cases was "leaving an account logged in" or "choosing a weak password" stated as a risk. Particularly weak passwords are a risk which security professionals and researchers have tried to get the general population to take seriously for a very long time. Unfortunately, participants did not consider these issues to be a major risk. I thus argue that the risk of choosing weak credentials and not logging out of accounts is either unknown to many users, is not important enough to be salient, or users are not aware that they are actually doing these and other security-relevant activities wrong.

6.4.5 The Way Forward

Based on the above results, I postulate that users are only aware of a limited set of risks without being prompted and this set includes many risks that are not addressed by security technologies. I found that the relative importance of risks when using the Internet is perceived to be low in the sample to start with and if a security mechanism only addresses a few of the risks users are aware of, the perceived relevance likely becomes even lower. I therefore find that, for a security measure to become relevant and be adopted, users need to be aware of a serious risk with personal and immediate consequences, which are addressed by the technology in question.

Alternatively, the results suggest that new security technology can be specifically designed to address the users' greatest existing concerns and therefore more readily find adoption. For example, it can be worth researching whether security measures protecting against man-in-the-middle attacks, such as visual indicators or warnings, might be more readily adopted if users were convinced that they prevent fraud and identity theft. Framing benefits around common scenarios and addressing the risks that are particularly salient in that scenario can help to tip the cost/benefit scales in favor of the security measure.

Another question that arises is whether or not it is possible to create additional awareness in users living in a modern society, with many other concerns competing for their attention. The results from Section 6.3.4 indicate that it may indeed be possible to raise awareness for particular risks. Malware is possibly the most common and long-standing security threat to end-user IT systems and the installation of anti-

virus protection is recommended to most PC users. However, before people started to use information technology on a daily basis, they probably weren't as aware of the malware risk, as the results of Friedman et al. from 2002 suggest [66]. Yet, users may have actually experienced malware on their own device or heard stories from friends or family about such events since then. Additionally, there is considerable advertising for malware protection products that also remind users of the risks and made them learn to be afraid of malware. It needs to be subject of future work to see if and how awareness can also be raised for other IT security risks.

6.4.6 Limitations

While two diverse user groups were sampled for this study, especially the incidence of individual risks cannot be generalized. I also aimed to make differences between the two chosen populations clear in the text but also admit that a complete picture can only be painted by redoing this study with a population representative of all Internet users. Similarly, I deliberately chose a particular set of scenarios to test the influence of context on risk awareness. Other scenarios will likely yield different sets of risks, for example when considering the use of different service providers, as a user may trust other services less. However, I argue that the patterns found in the results already hold valuable insights concerning the human aspects of IT security research. Future work is needed to look at effects a variation of trust in the scenarios may have. Additionally, I chose to use a survey as my research method, in order to obtain a wide view. It is possible that additional information can be obtained from using an in-depth interviewing technique, which is subject to future work.

Deliberately not providing a definition of the terms risk and consequence also has a potential influence on the presented results. I chose this approach to bias participants as little as possible. I argue that due to this choice, the results provided insights into how users define the concept of an IT security risk and its consequences for themselves in their everyday conduct. Forcing them to adhere to certain definitions could have led to results that are potentially closer to how security experts reason about risks and consequences, as participants may have provided additional risks and consequences they didn't think of or phrased those they provided differently. However, at the time users decide about whether or not to adopt a security measure, there also is no instruction sheet that provides a definition of how risk ought to be appraised before they make a decision. I am thus convinced that not providing definitions was a vital part for the investigation I presented.

Last, I used an inductive coding procedure to analyze the risks participants provided in open-ended responses. Analyzing the data showed that categorizing and coding risks and consequences is a task that can be tackled from many different angles. I adopted a pragmatic approach that allowed to get a general overview of risks. Yet, coding risks in different ways may allow researchers to gain additional insights into particular aspects of risk awareness.

6.5 Summary

In this survey of risk awareness during Internet use, I find that users only showed awareness of seven risks on average. While this was to be expected from non-experts, I posit that this also shows that the security community will have a hard time to get new security measures adopted in the general population under the premise that only risks users are aware of are considered in their adoption budget. Furthermore, I present evidence that the overall set of risks users perceive is very diverse and most of these risks were neither very specific nor can users easily protect themselves against these risks by using particular technologies. Additionally, existing security and privacy measures will often only address a small part of the risks users are aware of and focus strongly on the technical risks I found users are generally not too concerned about. This may then create the view that adopting a particular measure will not significantly reduce the risks of being on the Internet and is thus not worth the time, money and effort. I also posit that participants do often not see consequences that apply to them personally, effectively diminishing the benefits in terms of their adoption budget.

The main result presented in this chapter is that users are often not ready to invest effort into changing their behavior or adopting security measures for the above reasons. My analysis yields new insights into why certain security measures may not be adopted by end-users as well as which factors could influence adoption and hence need to be subject to further research. Security measures that aim to improve end-user security or privacy on the Internet would thus need to be designed to address salient risks and consequences as perceived by their users. The usable security community can support this process by further analyzing the protection needs of individuals, how security mechanisms can be tailored for adoption, as well as investigating possibilities to raise user awareness about important security risks, including their own negligence, effectively. The results presented in this chapter can serve as a foundation for this important field of future work.

7 The Way Forward – Improving Security Decision Dialogs Using Personalized Examples

After investigating human factors of the adoption of security measures as well as risk perception, I present an evaluation of a prototype that aims to improve adoption by leveraging personal examples for risk communication in this chapter. Users are reminded of the immediate and personal consequences of taking a security- and privacy-relevant decision during the installation process of smartphone apps. According to the results presented in the previous chapters, this should lessen the current ignorance of the security-relevant information presented during the installation process. I found that users made more security-conscious decisions when confronted with clear, understandable descriptions of risks as well as personal consequences in a lightweight and effortless manner.

7.1 Motivation

As detailed in Chapter 2, communicating the risk pertaining to certain actions is a long-standing problem in human-computer interactions and IT security in particular. Many IT security systems use a decision dialog to provide users with information about potential risks. However, recent research has repeatedly demonstrated that such dialogs are often ineffective and quickly ignored [21, 51] or provide information that is hard to understand for the user [77]. In contrast, Rader et al. found that informal stories influence security behavior and thinking as they are being relayed from one user to the next [132]. These stories offer concrete examples of good or bad things that happened to people which a user can relate to. Blackwell et al. [13] previously posited that abstract information in software causes a gap between system designers and users. Also, the results from Chapters 5 and 6 show that users are often not aware of concrete, personal consequences of the existing risks in IT systems. However, as I showed in Chapters 2 and 3, this is necessary for a positive appraisal of a threat which is a necessary prerequisite for adoption. In this chapter, I hence look at a common security and privacy decision many users face on a regular basis: granting permissions to smartphone apps.

As smartphones gain popularity, they also get more important in many people's daily life, managing a large variety of personal information, including emails, pictures, call logs, and text messages. On Android, this information is protected from unauthorized access using permissions. Users get to decide whether or not they

agree to the capabilities an app will have and which information on the phone will be accessible by that app after it has been installed. The most important prerequisite for such permission systems to be useful and secure in general is that the user can understand and decide which (sets of) permissions are okay to grant for an app and which might be harmful in a given context. Since apps can harvest and send out private information on the first launch, the current trial-and-error app installation behavior is critical from a security and privacy perspective. Previous research of Felt et al. has shown that only 17 % of Android smartphone users are consciously aware of the specific permissions an app demands during installation [60].

To overcome the abstract nature of the existing dialog, my modified version shows users personalized, concrete examples of capabilities the app would get and which information it could access, using personal information present on the phone. The goal is to raise awareness and make users more cautious while installing apps. For example, such a personalized example during app installation can say: "*If you install this app, it will be able to access and delete the following of your photos*", followed by a sample of the user's actual photos contained on the device. Thus, the communicated risks address a concrete, personal piece of information while listing a known cause as well as concrete consequences that a user can easily imagine. This can then allow users to judge whether or not a risk is acceptable or not: I may trust an anti-virus app to possibly delete some infected files in order to protect me from malware. Yet, I do not trust a random game to have access to text messages from my partner.

I emphasize that while users are probably aware of the private information they have on their devices and which apps they have installed, the individual relationship between an app's permissions and the concrete pieces of personal and private information that can then be accessed is never explicitly clarified. My novel approach aims to enable users to make informed decisions about the risks that apps pose to their private information by actually demonstrating which pieces of information can be accessed.

The following main contributions will be detailed in this chapter:

- I explore the effect of personalized security decision dialogs on the Android app installation process, leveraging the rich set of personal information available on smartphones.
- I show how to design decision dialogs with personal examples and present a prototype of this approach.
- I provide and discuss the results of two user studies, demonstrating the effectiveness of this approach compared to the standard dialog.

The results show that the modified dialog is able to cause a considerable amount of users to rethink an app selection because of risks arising from permissions, even after an installation decision has already been made. Additionally, users who saw the modified permissions dialog payed significantly more attention to an app's permissions in my experiments. I also found that a negative affect can increase attention.

In the following, I will outline additional related work, introduce and motivate the concept before presenting the results and a discussion.

Disclaimer: The contents of this chapter were previously published and presented at the Human Factors in Computing Systems conference (CHI) in 2014 under the title "Using Personal Examples to Improve Risk Communication for Security and Privacy Decisions" together with Markus Hettig, Susanne Weber, and Matthew Smith [78]. Susanne and Markus contributed to this work as part of their graduate coursework under my supervision. The co-authors assisted and advised me while developing the study design, helped with conducting the study, and provided feedback on the manuscript. An early version of this work was also presented at the Risk Perception Workshop of SOUPS 2013.

7.2 Related Work

On the theoretical side, the HITL framework [34] and the C-HIP model [162], which HITL is based on, are general information transmission models which describe the process of sending an information through a channel to the human receiver. After receiving the concrete information, attention must be shifted to this information before the comprehension process starts and eventually an appropriate reaction follows. It has been shown [10, 163] that a personalization of decision dialogs can support this process.

According to De Paula et al. [43], the central problem of human interaction with IT security systems is that users should be able to make informed decisions without further help. Additionally, Bravo-Lillo et al. [19] state that such dialogs should "inform clearly about the consequences of the actions" and about the risks involved and emphasize that it is important to inform about whether an action is safe or not right before the user ultimately decides. Bravo-Lillo et al. [20] also investigated the behavior of novice users confronted with situations in which they should make security decisions. It was shown that those users are not aware of the sensitivity of their data and mostly started to worry after deciding to allow access. They found that novice users often act with a "let's-see-what-happens" attitude without thinking about possible consequences.

Blackwell et al. [13] discussed the influence of abstract representations in computational tasks and suggests to center a user's understanding of a system around task completion to overcome abstractness. Similarly, Inglesant et al. [91] and Karat et al. [97] find that the use of natural language to create access control policies is beneficial for their quality. Also, Raja et al. [133] found that users have improved mental models of firewall operation when using metaphors to embed security decisions into an application context.

Egelman [51] recently investigated Facebook Connect security decision dialogs. The central finding was that habituation caused most users to ignore the content of the dialog, even though it was modified. Showing users personal information from their personal profiles, such as gender, name, or relationship status and sexual orientation,

which would actually be accessed by the third-party website, did not improve the dialog's efficacy in their study. In contrast to this work, the displayed information was already disclosed to at least one online service and the pieces of information cited in the dialogs were not very abstract or technical.

All of the above approaches argue for more concrete and graspable information in decision dialogs. Concrete examples of undisclosed private information have, however, not been used to highlight risks and possible consequences to support the decision demanded from the user. To the best of my knowledge, this is the first evaluation of personalized decision dialogs that leverage the large amount of personal information contained on smarpthones.

7.2.1 Permissions

A large body of work on Android permissions was compiled by Felt et al. (e. g., [57, 60]), investigating how permissions are used, how users perceive permissions, their attitudes towards potential risks as well as additional models applicable to ask for permission on smartphones. Most relevantly, they found that only very few users (3 %) actually understood which assets were protected by a given permission.

Pandita et al. [129] investigated to what extent a user's expectations of permissions match the actually requested set of permissions using natural language processing. They propose to automatically extract justifications for permissions from app descriptions and flag those apps where not all requested permissions are justified in the description. They postulate that this can help users to make informed decisions.

Recently, Kelley et al. [100] argued for the need of privacy as part of the app decision making process. They inserted an overview of the private information an app will be able to access into the overview page of Google's Play Store. Their results show that they were able to influence the user's decisions compared to the existing permissions display. Their goal was to make privacy part of the decision process by abstracting permissions into a summary table.

In contrast, I attempt to improve risk communication by making permission risks graspable and hence understandable. I want to enable users to take an ultimate decision if he or she is willing to trust a certain app and its developer with access to personal information. Using only the approach of Kelley et al., a user might just count the checkmarks in their Privacy Facts display and therefore make a privacy-aware choice. However, an app with only two permissions can already cause great harm if the app has malicious intent. I argue that more information is necessary to clarify the risks pertaining to certain permission sets than can be fitted on an overview page. The approach presented in this chapter is hence complementary to the Privacy Facts display of Kelley and colleagues.

7.3 Design

To leverage the power of personal examples as security decision dialogs for app installation, this information needs to be presented to the user in a concise and appealing fashion. Additionally, it is paramount that the permission dialog is able to make the user question an app selection that was already made. This is necessary since the permissions will be displayed on a separate page of the dialog, which only becomes visible after a button labeled "Install" has already been pressed and therefore a choice for an app has been made. In this section, I will present the design of the modified permission dialog for Google's Play Store and discuss its rationale. During the development of the UI, my co-authors and I ran several pilot studies to test prototype efficacy.

7.3.1 Permissions Visualization

The current permission dialog of Google's Play Store by default only shows a small number of the 79 permissions[1] which are deemed to be most important. The remaining permissions which an app requests can be displayed by unfolding a hidden panel. As Felt et al. already reported in 2011, the Android OS defines a large number of permissions of which many are rarely used [57]. Therefore, I wanted to choose a representative set of permissions for the evaluation.

To find the most common permissions, Markus Hettig crawled the 34,875 most popular apps on Google's Play Store in early 2013 and counted which permissions are requested. From the top 20 of the set of requested permissions, my co-authors and I then picked ten (see Table 7.1) that can affect private information.

Permission	Rank #	Requested By
full network access	1	82 %
modify external storage	3	56 %
read phone status and identity	5	42 %
view Wi-Fi connections	8	26 %
precise location	9	23 %
find accounts on the device	12	16 %
take pictures and videos	14	8 %
read contacts	15	7 %
read call log	17	6 %
retrieve running apps	18	6 %

Table 7.1: The permissions selected for the evaluation, their rank in the top 20 of permissions, and how many of the 34,875 crawled apps requested them.

To visualize each of these, several random examples are selected from the data that this permission allows access to and displayed alongside a concrete, one-sentence

[1] At the time of writing, Google did not specify the actual number. According to Au et al. [4], Android 4.0 offers 79 permissions that can be requested by a third-party app.

description mentioning the user's actual data. Choosing random examples for each permission can prevent habituation, as different content is visible each time the dialog is shown. Additionally, not every example drawn from the available information on the phone is considered equally private and displaying multiple examples shows the user a cross-section of the private information available on the phone.

Three of the permissions are notable exceptions: *full network access*, *take pictures and videos* and *location* do not allow access to existing data but can be used to exfiltrate information from the device or eavesdrop on the user's actions or surroundings. In these cases, I only used a description for the *full network access* permission or used the actual camera view for the *take pictures and videos* permission and the current location for the *location* permission.

I used the same general layout, fonts and headlines as in the existing permissions dialog for all visualizations to match the look-and-feel of the original store. I slightly increased the size of the headlines for each permission from 18 dp (density-independent pixels) to 20 dp to make them stand out more clearly from the examples themselves.

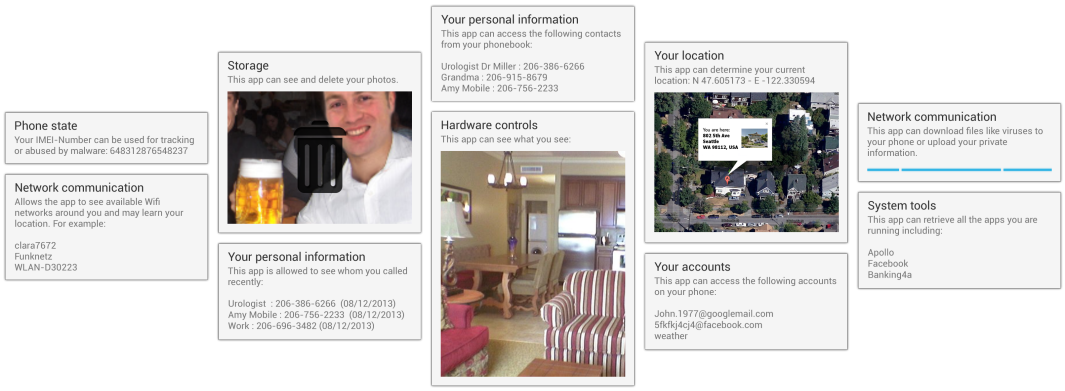


Figure 7.1: Overview of all permission visualizations created for the study.

Each of the permissions and its visualization are shown in Figure 7.1. I changed the descriptive text to mention real data and showed pieces of this data where possible. For example, for *full network access*, the user would be warned that this permission can be used to download malware to his or her phone or upload private information. Similarly, *modify external storage* showed a different picture from all the pictures taken with the phone and accessible via the storage permission each 1.5 seconds and stated that this app would be able to delete those. The phone's IMEI number and a statement that this number could be used for tracking or abused by malware were presented when the *read phone status and identity* permission was present. The remaining permissions gave examples of the information that could be accessed by this app alongside a statement saying "This app can see . . . " or "This app has access to . . . ". The display would then show a selection of three contacts, three previous calls, three accounts configured on the device, three currently running apps, three nearby Wi-Fi networks and the current location on a Google Maps satellite image. In each case, the text in the dialog explicitly mentioned that the app to be installed

will be able to interact with the concrete piece of data, in order to create a graspable connection between the app and the private information.

Overall, the permissions display was modified to include a different descriptive text more related to real data and a personal example of that personal data where possible. Additionally, the headline font size was slightly increased.

In a pilot study, my co-authors and I monitored participants' reactions to each of the permission visualizations and ordered them so that those deemed most relevant would be shown first. Contacts, call log and photos received the most attention and reactions from participants. More technical permissions such as access to the IMEI number and a list of Wi-Fi networks were therefore moved to the end of the list.

The remaining permissions were not part of the visualization but could be displayed as a fold-out panel at the bottom of the modified permissions display, similar to the existing approach. The choice of which permissions to display and which examples to use potentially influences the efficacy of the approach. I will discuss possible implications in the *General Discussion* section after presenting the results of the evaluation.

7.3.2 Play Store Integration

As already argued above, I chose to replace the existing Play Store permissions dialog, which acts as an ultimate decision before the actual installation. Figure 7.2 shows a comparison of the old and the modified permissions display. My co-author Markus Hettig created a working Android app simulating the existing Play Store, serving as a prototype for the evaluation.

7.4 Lab Study

After piloting the prototype with several users, I set up a lab study to evaluate the approach and gain some further insights into how users select apps and perceive the risks pertaining to an app's permissions. Susanne Weber helped me to conduct this study. The methodology is based on the work of Kelley et al. [100]. To improve ecological validity, I choose to let participants use their personal smartphones during the study and therefore base the evaluation on real private information.

7.4.1 Method

I invited owners of smartphones running Android 4.0 or newer from a university-wide study participation list to attend a lab study on app selection in the Google Play Store. The study design included limited deception by not mentioning the study's focus on permissions to prevent bias and I recruited only Android users to prevent side effects from unfamiliarity with smartphones, apps or the Play Store. The invitation offered 8 Euros of compensation for a 30 minute study and stated

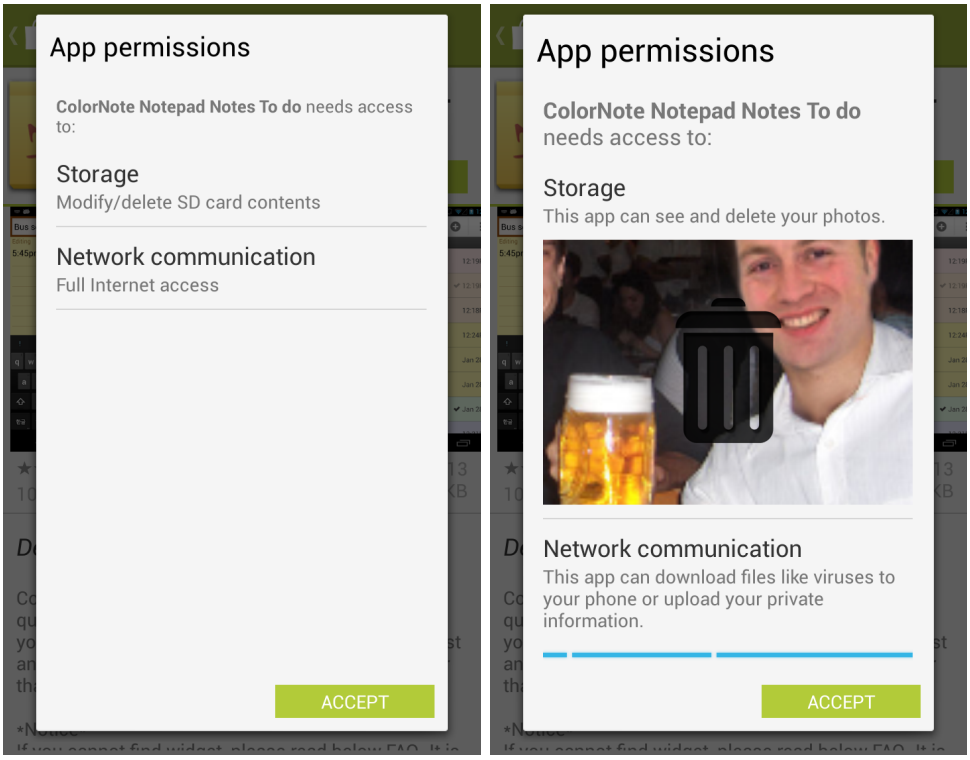


Figure 7.2: Comparison of the existing permissions dialog on the left and the modified version on the right.

that I would be installing a test app on users' personal device during the lab session. In the lab, each participant was introduced to the study and signed a consent form before installing the prototype app. Participants were informed that the app would be collecting usage data but no personal information and that I would transfer this information from their devices once the study was completed.

They were then instructed to complete six tasks, asking them to role-play the selection and installation of apps from a certain category that fit a certain purpose they were supposed to be in need of. Participants were also asked to think aloud while making their decisions. Within each category, participants ultimately had to choose between installing one of two apps with different sets of permissions (see below). They were also instructed that they may choose not to install any app in each category, if none of the available options suited their personal needs. I added the "none" option to cater for the fact that users can normally abort an app choice process at any time without installing an app. In terms of study design, both, my modified dialog as well as the approach of Kelley et al. have drawbacks. Adding the "none" option allows participants to ignore a difficult choice, while forcing a decision may not represent realistic behavior. After finishing the tasks, participants completed a questionnaire on general app installation behavior as well as their perception of permissions. At the end, I debriefed participants about the true purpose of the study, clarified that the Play Store app was only a mockup and no apps were

installed, removed all traces of the app from the participants' phones and gave them an opportunity to ask further questions.

The study used a between-subjects design with respect to the permissions dialog and I compensated for effects of fatigue and learning by randomizing the task order. The task order assignment was based on latin squares.

Test App To run this test, Markus Hettig implemented a mockup of the Google Play Store app as mentioned above. The mockup would either display the conventional representation of permissions as text or the modified representation using personal examples (cf. Figure 7.1). Before displaying the permissions, users were able to navigate through the Play Store as usual, compare several apps in the list view and look at app details, screenshots, and ratings. Each participant was provided with six tasks that asked them to find, select, and install an app with a certain purpose from a certain category of the Play Store. The mockup app measured the time participants spent looking at each app description and the respective permission screens.

To offer an experience as realistic as possible, the mockup store completely imitated the functionality of the real Play Store except for three differences. First, all app listings only included seven apps to not frustrate participants or have them spend too much time going through all available apps. Apps that did not fit the purpose of the task were not selectable. Second, there also was no search function, since the pilot study showed that participants would get frustrated searching for apps they already knew but were not included in this test. Instead, participants were instructed to navigate using the categories. Third, the apps that suited the respective tasks would be randomly displayed as first or second item in the "Top Free" list to facilitate discovery and again make the tasks less frustrating without creating a bias for one app due to its position.

Apps Each category contained two apps that actually fit the given purpose for the respective task. Within each category, padding apps were included to create a more realistic Play Store mockup. All apps and their names, descriptions, screenshots, permissions, and ratings were taken from the real Play Store and the 12 relevant apps (cf. Table 7.2) were selected to have similar functionality, average rating, and visual appeal. Similar to Kelley et al. [100], I also chose to display one 2- or 3-star rating, one 4-star and one 5-star rating for each app. I selected apps that were likely to be unknown to my participants. I also included two app categories to test additional factors, rating and brand, to allow for assessing the influence of ratings and brand recognition on risk perception. Therefore, one photo app had a medium rating of 3.4 and another a high rating of 4.7. In another category, the well-known *Google Search* app was available with the *Quick Search* app. Concerning other properties the apps were again as similar as possible.

I also largely preserved the existing permissions for the apps. Yet, a larger variety of permission differences between the apps was desirable to see if there is a threshold

of difference in terms of permission sets that is necessary for my approach to work. I therefore added or removed one or two permissions in four apps (cf. Table 7.2). Most notably, I made the Tetrity app not request any permissions and PicsArt, in addition to having a high rating, request all but one permission.

App Name	Network Access	Ext. Storage	Phone State	WiFi Connections	Location	Find Accounts	Take Pictures	Read Contacts	Read Call Log	Running Apps	**Total**
EasyMoney		•	•				•				**3**
CWMoney	•	•	•		•		•	○			**6**
ColorNote	•	•									**2**
CatchNotes	•	•	•		•	•	•	•	•		**8**
Tetrity	⊖		⊖								**0**
Traris Deluxe	•		•	•	•						**4**
Weather	•			⊖	•						**2**
Eye in the Sky	•	•		•	•						**4**
PhotoEffects	•	•		•							**3**
PicsArt (rating)	•	•	•	•	•	•	•	○		•	**9**
Quick Search	•			•							**2**
Google (brand)	•	•		•	•	•		•	•		**7**

Table 7.2: The apps used in the studies and their respective permissions. A ○ indicates an added permission and a ⊖ indicates a removed permission. Two apps were part of a particular category (upper apps always requested less permissions) and participants were asked to choose one of those or none.

7.4.2 Participants and Results

After pilot testing the experimental setup, $n = 36$ participants completed the study. Sessions lasted between 15 and 30 minutes including the task, the questionnaire and the debriefing. Participant demographics can be found in Table 7.3. For the Westin index, I used the most recent set of question from Westin's 2001 Internet Privacy survey [108]. I did not find any significant differences in the measured data based on these demographic properties, including privacy inclination.

The left hand side of Table 7.4 details the results of the lab experiment, providing installation counts for each app. Across all six choices, more participants opted to install no app of the available two with the modified permissions dialog when compared to the existing Android permissions. Yet, as shown in Table 7.4, the effect is only significant in half of the app choices. Also, in four cases participants tended to install the less-requesting app or no app rather than the over-requesting app. The well-known brand and higher rating of two of the over-requesting apps

N	36	
age	19 – 30 years	
	median 23 years	
gender	12	female
	23	male
	1	N/A
IT experience	7	with professional or educational IT experience
	2	students of computer science
Westin Index	18	privacy fundamentalists
	17	privacy pragmatists
	1	privacy unconcerned
Incidents	17	previously victims of online dangers
	3	unsure

Table 7.3: Participants demographics for the lab study.

did not diminish this effect. Similarly, I did not observe any effects for the Tetrity app, requesting no permissions. The PicsArt app, however, requesting almost all permissions, had the greatest reduction of installation counts (yet not quite significant). Participants became aware of the large number of permissions requested, even though this app's rating was considerably higher than for the alternative. This is also a notable difference to the results obtained using the Privacy Facts display of Kelley et al., where the rating was actually more important than the privacy facts.

Participants only spent an average of 3.1 seconds (median 1.0s) looking at the old permissions and 7.6 seconds (median 2.4s) on the modified version. While these values differ significantly between the two permission displays (repeated measures ANOVA across apps, permission dialog as between-subjects factor, omnibus $F(1, 34) = 4.98$, two-tailed $p = .03$), the time spent on each permission dialog is still very brief.

Installation Behavior I also asked participants about their installation behavior with free apps. 19 (52.8 %) stated that they usually look at several apps and then install one of them. An additional 11 (30.6 %) said they would install multiple apps, try them and then uninstall apps that did not suit their needs. 5 participants said that they would use a mix of the previous strategies. One participant stated to just install a number of apps and trying them without paying much attention to any descriptions. I then asked participants to rate how similar their usual selection behavior is to their behavior in the study. All but four participants selected 5 or more on a scale from *not similar at all* (1) to *very similar* (7). This suggests that the observed app choices are a suitable approximation of real behavior.

	Lab Study				**Online Study**			
App Name	Existing Permissions Dialog (n=18)	Personalized Permissions Dialog (n=18)	Δ_l	p-value	Existing Permissions Dialog (n=157)	Personalized Permissions Dialog (n=175)	Δ_o	p-value
EasyMoney	77.8 %	33.3 %	-44.5 %		36.3 %	33.7 %	-2.6 %	
CW Money	5.6 %	22.2 %	+16.6 %		49.7 %	28.6 %	**-21.1 %**	
none	16.7 %	44.4 %	+27.7 %	**.028**	14.0 %	37.7 %	**+23.7 %**	< **.0001**
ColorNotes	88.9 %	50.0 %	-38.9 %		79.0 %	57.7 %	-21.3 %	
CatchNotes	11.1 %	11.1 %	± 0 %		12.7 %	13.1 %	+.4 %	
none	0.0 %	38.9 %	+38.9 %	**.006**	8.3 %	29.1 %	**+20.8 %**	< **.0001**
Tetrity	61.1 %	55.6 %	-5.5 %		49.1 %	68.0 %	+18.9 %	
Traris Deluxe	38.9 %	27.8 %	-11.1 %		44.6 %	22.9 %	**-21.7 %**	
none	0.0 %	16.7 %	+16.7 %	.28	6.4 %	9.1 %	+2.7 %	**.0002**
Weather	66.7 %	77.8 %	+11.1 %		75.8 %	70.9 %	-4.9 %	
Eye in the Sky	27.8 %	5.6 %	-22.2 %		17.8 %	15.4 %	-2.4 %	
none	5.6 %	16.7 %	+11.1 %	.2	6.4 %	13.7 %	+7.3 %	.081
PhotoEffects	16.7 %	33.3 %	+16.6 %		19.8 %	28.0 %	+8.2 %	
PicsArt (rating)	83.3 %	50.0 %	-33.3 %		71.3 %	49.1 %	-22.2 %	
none	0.0 %	16.7 %	+16.7 %	.068	8.9 %	22.9 %	**+14.0** %	< **.0001**
Quick Search	27.8 %	27.8 %	±0.0 %		15.9 %	20.6 %	+4.7 %	
Google (brand)	72.2 %	44.4 %	-27.8 %		83.4 %	64.6 %	-18.8 %	
none	0.0 %	27.8 %	+27.8 %	**.041**	.6 %	14.9 %	**+14.3 %**	<**.0001**

Table 7.4: Installation count results of both, lab and online study. The table shows the percentage of participants that chose to install either one of two apps from each category or none, with the bottom app requesting more and unnecessary permissions than the top app within each group of two. Bold typesetting indicates significant contributions to the χ^2 value (standardized residual > 1.96) and significant p-values according to Fisher's exact test on the respective 3 (choice) $\times$ 2 (type of permission dialog) cross-table.

I also asked participants to state how frequently they have not installed an app because of several factors (cf. Figure 7.3). Most participants said to not have installed an app because of the rating or the cost of an app, while the number of requested permissions or improper permissions only caused about a third of the participants to not install an app more than five times. This underlines the small amount of time users spend viewing the permission display. Additionally, most participants did not see many risks arising from malicious apps in the Play Store: they gave an average rating of 3.4 (median of 3) on a scale from "very low risk" (1) to "very high risk" (7).

However, there was a slightly higher concern for the amount of danger for private information arising from smartphone apps in general: participants gave a mean and median rating of 5.0 on a scale from "no danger at all" (1) to "great danger" (7).

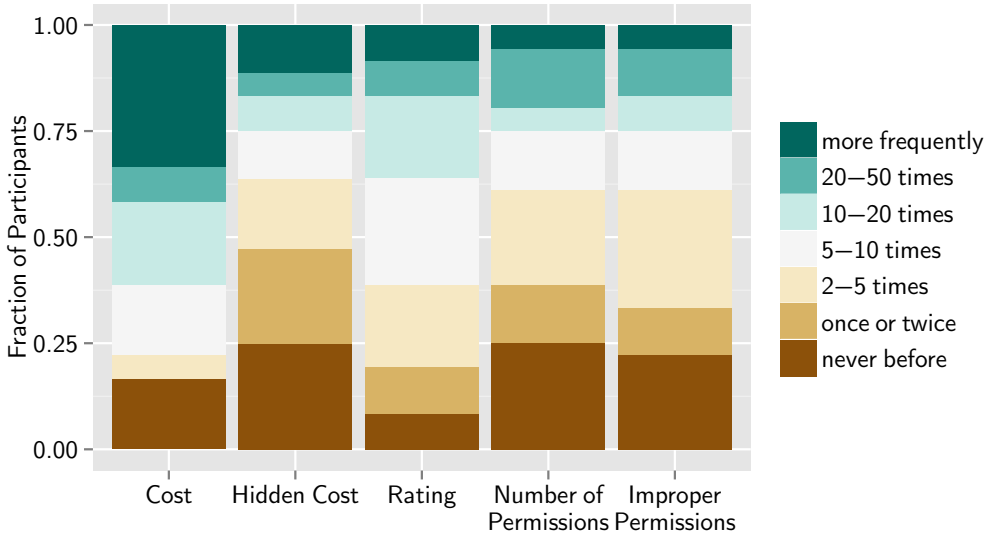


Figure 7.3: Frequency of not installing an app because of different factors from the lab study.

Qualitative Insights Relevant comments users gave while thinking aloud during the app selection tasks were collected by the interviewer. Most interestingly, many participants indicated to trust Google to curate the Play Store or to trust the community of Android users to flag malicious apps. This confirms the slightly better rating for risks arising from Play Store apps compared to apps in general presented in the previous section. Furthermore, several participants justified simply accepting any set of permissions without checking based on this view. Similarly, many participants stated to have "*nothing to hide*" or to be "*too uninteresting*" for anyone to attack them. Other participants also said that they do not put sensitive private information on their smartphone, because "*Google can see everything anyway*" or because existing protection mechanisms "*do not pose serious problems for hackers*".

The interviewer queried those participants who had seen the new permission dialog about which of the private information contained in the dialog was most sensitive. Those participants mostly referred to the pictures as being undesirable for others to have access to but also found that phone book, call log, and location can be sensitive. One participant said "*there are contacts in my phone book for which I would not like others to know that I am in contact with those*". Several participants hinted at a negative affect ("*I felt very alarmed which data can be accessed by apps*") or were very surprised how far access can go ("*I would not have thought that it can see or even delete all these things!*").

7.4.3 Discussion

The lab study yielded encouraging results. The modified permission dialog was able to generate a significant effect on the participants' installation behavior. The participants' comments support that seeing examples of their personal data encouraged reflection about possible consequences and whether or not they are ready to trust the developer and accept the risks. However, participants' comments and questionnaire responses also indicate that many other factors compete with permissions for the installation decision. A certain amount of the variation found in the results presented above is therefore accountable to users' different tastes in apps, as the baseline installs show. Several participants also stated to completely ignore permissions, because they don't feel that their personal data is threatened by apps or that they have nothing to hide. It was not the aim of this study to make these users mind their security and privacy if they don't have an interest in it. However, it is conceivable that making the risks more concrete can influence some users' views of apps' security and privacy impact.

The users' choices during the study also showed that, while I chose the four sets of apps with equal functionality, rating and potential for visual appeal, some apps were already clearly preferred in the baseline condition. However, in the cases where the over-requesting app was preferred, the permission dialog still had an impact and was able to make users rethink their choice (cf., for example, the Weather or ColorNotes app in Table 7.4). Additionally, the effect I found on the rating-category underlines that the modified dialog was indeed able to overlay decisions already made, even when some factors are strongly in favor of the more dangerous option.

Furthermore, the examples of private information displayed to communicate the risks of installing a particular app appeared to be making many participants understand the extent of what a permission allows an app to do. Even though all participants had been using Android smartphones for several months and had installed several new apps, seeing the permissions displayed in this new fashion gave them a better idea of what permissions really entail. Participants that came in contact with the new permission dialog stated that they would "*pay more attention to these things in the future*".

7.5 Online Study

To confirm the effects observed in the lab experiment on a limited population, I went on to evaluate my approach with a more diverse population. Since the lab study showed large individual differences in the way apps are selected and installed, a larger sample would be required to obtain more reliable results. Also, students are often considered to interact differently with technology than the general population and may have therefore biased the lab results. Additionally, I wanted to investigate to what extent this novel approach caused participants to be afraid of misuse of their private information. Such a negative affect could be a key motivation for less risky installation decisions and influencing factor for threat appraisal. The questionnaire in this study therefore included questions that elicited how afraid of misuse of private information participants felt as well as how aware participants were of an app's control over personal data through the use of the permission dialog.

7.5.1 Method

I planned the online study to largely resemble the lab study. To access a diverse population, I decided to use Amazon's Mechanical Turk service to run the experiment. While MTurk also does not provide a sample that is representative of the general population, this service is commonly used to access a population that is more diverse than usually available. The task was framed as the common role-play of helping a good friend (previously applied in e. g. [20, 100]), who just bought a smartphone and does not know which apps to install. In the study, participants would be shown a description of the scenario, including a picture of the friend John and his family to create a feeling of familiarity. Participants were asked to imagine that John gave them his phone and a list of activities he could use an app for. They were presented with two suitable apps for each of the six activities. The order of the app choices was again randomized to control for effects of learning and fatigue. After finishing the selection process, participants completed a questionnaire as in the lab study and were compensated with $ 1.50. Only participants regularly using a phone running on Android 4.0 or newer in their daily life were asked to participate, in order to have similar levels of familiarity and habituation with smartphones and app installations. Participants were asked to specify their Android version in the questionnaire and I excluded participants who specified a lower version.

As Amazon's terms of service do not allow to ask MTurk workers to install applications, the tasks needed to be recreated in the browser. Markus Hettig and I created a web experiment where participants were presented with two side-by-side screenshots of the same apps for each of the same six tasks as in the lab study. Again, I chose a setup based on the work of Kelley et al. [100]. They were then asked to decide whether they would want to install any of the two apps. Again, they had the option to install none of the two. If participants chose to install an app, they were presented with a screenshot of the respective app's permission dialog (cf. Figure 7.2 and Table 7.2) and asked whether they want to continue with the installation or

return to the app overview. Half of the participants were randomly assigned to the modified permission dialog, which contained artificial private information of John (cf. Figures 7.1 and 7.2). After each choice, participants were asked to explain the reasons for their selection.

7.5.2 Participants and Results

332 MTurk workers successfully and validly completed the task. I included attention check questions in the questionnaire and the tasks, which I used to exclude 49 participants answering inconsistently or specifying an Android version less than 4.0. Table 7.5 provides an overview of the online study participants' demographics. I found no differences in the measured variables with respect to these properties, except that participants indicating previous professional IT experience chose the Google app significantly more frequently (Fisher's exact test, $p = .035$, odds ratio .51).

N	332	
Age	18 – 64 years	
	median 27 years	
Gender	38.6 %	female
	61.4 %	male
Occupation	50.6 %	full-time employees
	19.3 %	students
	10.2 %	part-time workers
	7.9 %	self-employed
	7.0 %	Homemaker
	4.3 %	Unemployed or retired
IT Experience	27.1 %	have worked in or studied IT
Smartphone Use	18	months (median)
Westin Index	29.8 %	privacy fundamentalists
	52.1 %	privacy pragmatists
	18.1 %	privacy unconcerned
Incidents	38.9 %	previously victims of online dangers
	6.3 %	unsure

Table 7.5: Participants demographics for the online study.

During the app choice tasks, participants again were significantly less likely to install the over-requesting apps when using the permission dialog with personalized examples as opposed to the baseline regular Android dialog. The right hand side of Table 7.4 gives an overview of the results in comparison with the results from the lab. Again, brand and rating did not diminish the effect. The online study hence confirms the effect found in the lab study.

The Weather app was again so popular in the baseline condition, that I did not find significant effects in this case. However, this also suggests that the modified dialog did not blindly scare users into not installing any app at all: the permission set of

the Weather app was apparently reasonable enough to be accepted for its purpose. This is also mirrored in the participants' reasoning about this choice: "*The only permission it needed was my location. That makes sense.*" "*It need [sic] permissions which I would expect from this kind of app*". In other tasks, users responded differently: "*Both apps had permissions that I wasn't comfortable accepting*". Furthermore, I found that 88.6 % of online participants with the new permission dialog installed three or more apps each.

Similar to the lab results, the permission dialog had a significant overall effect on the time spent viewing the permissions (repeated-measures ANOVA across apps, permissions dialog as between-subjects factor, $F(1, 330) = 53.98$, $p < .001$). Yet, Bonferroni-Holm-corrected pairwise comparisons only yielded significant results for the PicsArt and Quick Search Widget apps. However, the time spent looking at permissions was still short with 5.6 seconds (median 2.5s) in the baseline condition and 8.5 seconds (median 5.3s) in the modified condition. It is important to note that these values contain network and rendering delays, as they were collected on the server-side.

Again, I asked participants about their installation behavior with free apps. Similar to the lab, 208 (62.7 %) stated to examine several and then install one, 88 (26.5 %) try multiple apps and then uninstall those that did not suit their needs, and 31 (9.3 %) just try several apps without paying much attention to descriptions. 5 participants (1.5 %) said that they would use a mix of the previous strategies, while also considering external information sources as well as the required permissions. 89.4 % of participants again indicated to have behaved similarly in the online study, answering with 5 or more on a scale from *no similarity* (1) to *great similarity* (7). Online participants were also slightly more concerned about malicious apps in Google's Play Store (mean rating 3.96, median rating 4; 7=highest concern, see lab results) but slightly less concerned about the amount of danger for private information arising from smartphone apps (mean rating 4.2, median rating 4; 7=great danger).

Concerning participants' awareness of an app's control over personal data after completing the task, I found that the modified permissions dialog had a significant effect: more than twice as many participants compared to the existing dialog (87 vs. 40) chose the highest rating on a scale from *not aware* (1) to *very aware* (7) (odds ratio 2.88, Fisher's exact test, $p < .0001$). To assess how afraid people felt after seeing permissions, I asked them to rate to what extend the display of app permissions caused them to be afraid of John losing personal data or information. Similar to the general awareness, four times as many participants (13 vs. 57) indicated the highest rating on the provided scale from *not afraid at all* (1) to *very afraid* (7) (odds ratio 5.32, Fisher's exact test, $p < .0001$).

Most interestingly, I also found a significant effect of the above rating on the time spent viewing the permissions (repeated-measures ANOVA across apps, permission dialog and highest fear rating as between-subjects factor, $F(1, 328) = 124.3$, $p < .0001$). Participants that were very afraid to have John's private information compromised spent 5.4 seconds longer looking at both versions of the permission dialog. This is a twofold increase. Additionally, there was no significant interac-

tion between the two between-subjects variables, such that I did not find differences in being afraid and paying more attention with regard to which permission dialog participants saw ($F(1, 328) = .007$, $p = .8$).

7.5.3 Discussion

Even though the setup of the online study differed from the lab study and did not allow for using participants' actual private information to personalize the permission dialog, the results show that the approach was still effective. Participants were more likely to choose an installation option that requested less permissions. Furthermore, the Weather app showed that reasonable sets of permissions were recognized and accepted if they fit the app's purpose. The online study also suggests that communicating risk to users with examples created more awareness in participants and instilled a negative affect which caused them to pay more attention to the permissions. I therefore posit that this approach to risk communication can more easily override choices already made than existing approaches, especially when important information on the risk can only be presented afterwards.

7.6 General Discussion

The two studies have shown that personalized, more concrete examples and descriptions can be leveraged in security decision dialogs to increase their efficacy. The modified dialog was able to significantly impact the choice of apps in the Google Play Store, given that some apps were a greater risk to participants' security and privacy than others. The results also demonstrate that using personal examples in these dialogs can leverage users' affect to increase attention. While the modified dialog may have caused some users to shy away from installing any apps in the study, it will also make them consider the security-tradeoff they are about to enter as it actually is. Seeing the personal examples might also be scaring users into choosing conservatively, but users were not completely oblivious to what happens, as in many cases, an app was still installed. Users that did not install an app during the one try they had in the study would probably try to find another one at a later point or when they have more options as they still have a need for the desired functionality.

For this exploration, I used a mix of drastic and more neutral scenarios in the permissions display. It is, for instance, rather unlikely that an app will simply delete images while it is more plausible for an app to access and possibly upload contacts from the address book. Since it is hard and computationally expensive to automatically determine what actions an app actually can or does take, it is up to future work to determine which intensity of descriptions and which visualizations can work best in which situations. I posit that a balance between intensity and frequency of displaying the dialog needs to be found, since it is likely that if the most extreme examples are routinely used, they will lead to a fatigue effect if the

examples shown never happen. This is of course a fundamental problem of all risk communication.

The concrete choice of examples is also likely to be an additional factor for the efficacy of the risk-communication-by-example approach. Displaying photos of one's cat or a random landscape from some city the person has visited probably causes less of an emotional reaction than a portrait of one's partner or child. Since I only displayed random examples in this initial exploration, selecting specific and more private examples can potentially further increase the efficacy of personalized security decision dialogs.

Furthermore, the comments users gave during the experiments indicated that using the prototype only for a short amount of time already caused them to consider changing their general attitude towards permissions ("*I think I will pay more attention to those permissions in the future.*"). Using personalized examples may therefore also be a suitable tool for an initial or recurring education campaign in systems that can serve to make users understand the risks present in this system.

The prototype and scenario used in the evaluation assume a single-user environment for a smartphone. While I find this to be reasonable in many cases, personalized decision dialogs may also be desirable in situations where a device is shared by multiple users. While a shared device, such as a tablet, will probably yield less information that is especially private to one of its users, a system would need to make sure that the personalized dialogs do not create a privacy issue themselves by disclosing private information to others. For mobile devices, this is an important and unsolved problem concerning many other aspects of the app model as well and is the subject of ongoing work (e. g., [82]).

Similarly, app installation may take place in public and hence cause privacy issues by displaying private information in the permission dialog. To avoid this, users who frequently install apps in public should be able to opt out of using personal information or have the information displayed only after pressing a button.

7.6.1 Limitations

The study presented in this chapter is a first contact study and thus the novelty of the dialog itself may have increased the effects I observed. Additionally, I made the design decision to slightly increase headline size to separate the examples better, as well as to change the descriptive text of each permission to be less abstract and better fit the personalized context. These two changes may have confounded the results even though participants' qualitative reactions mainly concerned the personal information contained in the dialogs. Additionally, while I attempted to choose a set of apps which were not too well known, users may have already been familiar with some of the apps in the experiments, influencing their installation decision.

As stated above, habituation can also occur with a personalized permission dialog. However, the displayed examples are chosen randomly and hence the dialog changes between each app installation and therefore at least has a chance of countering

habituation, especially since the users regularly create more personal data (e. g. new photos, contacts, and call log entries). This naturally needs to be investigated with a long term study before any reliable statements can be made.

7.7 Summary

In this chapter, I demonstrated the value of minding human factors when designing a security measure. Leveraging personalized examples to improve risk communication for security and privacy decisions improved users' choices during the Android app installation process. Two experiments with diverse populations have shown that users make more risk-aware app choices when presented with concrete examples of the information at risk from undesirable permissions. I find that my approach has the ability to let users rethink a choice that they had already made, which was previously thought to be very difficult by many researchers. Also, improving the existing permission dialog in its current form has the advantage that no information is lost – as is the case in the privacy facts display of Kelley et al. – so that users can still detect malicious intent even though only few permissions may be requested. Additionally, displaying changing examples also has the potential to overcome the negative effects of habituation in a natural way.

The previous chapters of my thesis have repeatedly outlined that it is of central importance to provide graspable information in a concise way while not creating an additional burden for the user. The concept of personalized examples in security decision dialogs proposed in this chapter serves as a perfect demonstrator of the insights my thesis provided. The work presented in this chapter demonstrates that following the results I detail in the previous chapters can lead to improved adoption of security measures, as a large amount of users began to mind the permission dialog. Showing users how they are personally affected by their choice in the decision dialog made them pay more attention while effortlessly integrating this additional measure into their existing workflow.

Finally, the results I presented in this chapter provide some evidence that when security decision dialogs display personal information, a negative affect is likely created in users which increased their attention. This is an important direction for future work, as the role of users' affective states and emotions towards security measures and the risks these measures protect against have not been investigated in detail yet. I posit that leveraging emotions is a powerful way to make information graspable for users and also empowers them to engage with the information presented as a basis for their decisions. In terms of their adoption budget, risks that repeatedly elicit strong emotions may also be able to increase users' long-term awareness of certain risks and therefore tip the scales in favor of protection measures.

8 Conclusion

> *I don't pretend we have all the answers. But the questions are certainly worth thinking about.*
>
> — *Arthur C. Clarke*

Throughout this thesis, I investigated the influence of human factors on adoption decisions for IT security measures from a diverse set of angles. First and as a exemplary introduction into the topic at hand, I detailed the reasons for the lack of adoption of electronic identity features contained on the new German identity card (nPA). During this case study, it became evident that the ecosystem is important and should not be neglected when designing and deploying security measures. In the case of the nPA, a lack of suitable services which users gain an added value from as well as an uncomfortable and costly way of interacting with the system lowered adoption potential severely. This case study shows that new security technologies need to provide obvious utility and reduce the actual and usability costs of applying the system. Furthermore, social sources of information should be considered and engaged when trying to "advertise" a new technology. The eID system in the German identity card was criticized by security experts immediately after its launch and this bad press has stigmatized the system ever since.

Second, I presented a study that drilled down further on the impact of user interface usability for adoption intention of a security measure. The results showed that adding an indirection to the workflow, causing additional effort by requiring keys to be exchanged or additional buttons to be pressed, lowered adoption intention. This finding also closely relates to the results of the study in the eID chapter: having to purchase a card reader and needing this device to use the system severely impacted participants view of system utility. Furthermore, I found that making the security entirely invisible diminished participants trust in the system's security. This provides evidence on the need of reassurance and transparency about security properties for the user. Finally, many participants stated that they would not adopt a mechanism that does not offer key recovery. Altogether, this chapter showed that security measures need to impact the users' known workflows as little as possible but should not be fully invisible in order to be adopted.

Third, the study on risk perception and behavior of smartphone lock screens investigated the impact of self-efficacy and context on the adoption of a security measure. The central takeaway of this study is that measures are only perceived as useful if they provide protection beyond existing precautions. In particular, this study described a plethora of physical and manual ways of protection users routinely applied to divert harm from their smartphone. These measures were also context-sensitive

and the study was able to show that the perceived need for protection also varied between private, semi-public and public contexts. The added protection a lock screen provides was hence not always welcomed by users and their justifications and situative adaption of their behavior were able to explain their views. The study also demonstrated that the risk of unwanted access to the phone is not the most important concern for users. For many participants in this study, the monetary value of the phone itself was most important and trying not to lose or damage the phone was thus a top priority. Finally, the collection of usage data revealed that unlocking a smartphone can consume a considerable amount of the overall usage time. Participants spent up to 9 % on average of their smartphone usage time interacting with the protection mechanisms. If improved protection measures consume even more time per unlock attempt, they will likely be rejected by participants. This work again confirms the previous findings, that a security measure needs to be effortless and provide tangible benefits to justify their adoption.

Fourth, I presented a study on risk perception during daily Internet use. In the previous chapters, I repeatedly argued for the explicit need for tangible benefits in new security measures. Existing theories have also proposed that the adoption of security measures in organizational contexts is a rational decision of users, weighing costs in terms of effort and benefits in terms of avoiding punishment from management. Transferring this reasoning to a very common task, browsing the Internet at home, I investigated if the foundations necessary to be able to see a benefit exist in users. I posit that if users do not perceive risks and associated consequences that they want to be protected from, a security measure cannot generate a perceived benefit that justifies the effort to be invested in adopting a security measure. Generally this would mean that, again, only effortless security measures stand a chance of being adopted. The findings of my study support this view, as participants were only aware of very little concrete and personal risks and consequences. Additionally, the majority of risks was not addressable by IT security technology but were rather more or less inherent to using the Internet per se. Also, this study confirmed that risk perception and therefore readiness to adopt a security measure is dependent on usage context. Security measures that are intended to be widely used will hence need to provide a diverse set of benefits that addresses the perceived risks in many common tasks.

Last, the insights of the previous chapters were translated into a prototype of an improved security measure for the display of app permissions during the installation process on Android smartphones. The improved measure was designed to be effortless for the user, as it addresses a risk many users are currently not aware of. It also attempted to communicate a graspable, concrete and personal risk and the associated consequences in order to create an opportunity for a perceived benefit. Leveraging the rich set of personal information on the smartphone, the personal examples of risks and consequences were able to convince a significant amount of users to make more secure app installation decisions, i. e. this measure was adopted by users and helped them to secure their phone and private information. This prototype was hence able to show that following the results presented in this thesis

can increase the chances of adoption for security measures and therefore make users more secure.

8.1 Outlook

Working on issues influencing the adoption of security measures for several years, I can mainly see two ways for improving end-user security in the future: The easier one is to directly address risks which users are already aware of, which are salient to them and which they also currently aren't well protected against. Designing security measures around these risks will make it easy for users to perceive a benefit and hence adopt the measure. Alternatively, security measures need to be entirely effortless, so that there is no need for a perceived benefit. The second way is trying to support the process of changing users' risk perception, for example using risk communication and education.

Additionally, my work showed that personalization of security measures is a promising approach for future work, as it makes risks and consequences more accessible to the user which will hence influence the decision for or against adoption of this measure. For the prototype presented in Chapter 7, additional work could shed light on long-term effects of this approach, especially with respect to habituation and general changes in behavior with respect to privacy. As the presented prototype uses a random selection of personalized information, it has the potential to counter habituation effects. Security decision dialogs with personalized examples can also be a valuable tool for other scenarios, including SSL warning messages, software installation on desktop computers or posting on social network sites. I also posit that personalized decision dialogs can serve as an educational tool, that may also work retrospectively. A user could be confronted with examples of private information – such as sent or received emails – that have been transmitted without proper privacy protection or encryption to see whether or not this causes more privacy risk awareness or a demand for better security measures.

Creating a demand for security measures is also worth looking into. If users were feeling a need for better protection through security measures they would also be ready to invest more effort. However, it remains unclear how such a demand would be created. Existing security incidents rarely impact an individual's life and therefore a lack of IT security measures rarely has severe consequences for any single user. I posit that until a large part of the population has suffered from an incident that could have been prevented by appropriate IT security measures, such a demand is impossible to create. As long as security breaches only cause non-material damage, such as leaked email addresses and passwords or a loss of privacy, the potential benefits of IT security measures will remain low.

The evaluation of the Android permissions prototype also revealed that emotions may be playing a neglected role for the adoption of security measures. The study found that users feeling a negative affect were paying more attention to the provided information and that showing users personal examples was also able to create such

an affect in a greater number of users. If future security measures were able to trigger such emotions reliably, users' attention to the important information in security decision dialogs could be heightened. Emotions may also be able to change risk perception and therefore make users accept to invest more effort into the use of certain security measures. Additional studies are necessary to investigate the relationship between emotions and the adoption of security measures.

A Appendix: eID Focus Groups

A.1 Question Plan

Note: This question plan uses colloquial language on purpose to create a comfortable athmosphere. It was also translated from German for inclusion in this thesis.

Introduction Good morning and welcome to this group discussion. Thank you very much for taking the time to participate in this discussion about user behavior on the Internet. My name is Marian Harbach, I am a PhD student with the Distributed Computing and Security Group, here at the faculty of Electrical Engineering and Computer Science. Over there is Sascha Fahl, a colleague of mine, and he is going to help me with conducting this group today.

At the moment, we are working towards a better understanding of how users perceive the Internet and how they apply security protection mechanisms online. Through our work, we hope to find better and more secure solutions for end-user security on the Internet. Today, we would like to hear your views about the authentication process of websites, meaning the registration and log in process.

As students, you were invited to participate in this group since you rely on the Internet on a daily basis, especially for your work and personal communication. You also frequently use websites that require you to register and log in. We want to benefit from your experiences.

A few words about this group discussion: First of all, there are no right or wrong answers. We even expect you to have diverging views and we therefore want to encourage you to express views that are different from what other participants have previously said. This is also why we are recording this session, so that we do not miss any of your comments. We will of course make sure that all data is only analyzed in an anonymized fashion.

You can also see that there are name tents in front of all of us. Those help me to remember your names, but they can also help you. During the discussion, you do not need to address your responses and view towards me. If you are responding to something that one of your fellow students just said, of you want to express agreement or a contrasting argument, or if you want to add an example, please do this directly towards the previous speaker. You can basically ignore me and just have conversations amongst each other.

I am only here to pose some questions from time to time and listen. I also want to ensure that each of you has the chance to express their view, as we are interesting in

the experiences of every single one of you. So, if one you is talking a bit to much, I might ask this person to let others also have a chance to say something. And if one you is not saying much at all, I might ask this person directly to share their view about the topic at hand. We just want to be sure to capture all of your views and ideas and that you have an equal chance of sharing these with the group.

If you brought your mobile phone today and haven't put it on vibrate yet, please do so now. If you need to answer a call during the discussion, please feel free to do so outside. There's also a table with refreshments, coffee, and some sweets over there. Help yourselves at any time.

Okay, let's begin now. First of all, let us find out a bit more about each other by introducing ourselves briefly one after the other.

Opening Questions

1. To introduce yourselves, please briefly state your name and what are the most useful sites and services for you on the Internet.

Introductory Questions

2. For many sites, one needs to register a username and password and than log in using those. What are your experiences with these things?

Transition Questions

3. I brought a list with several categories of websites. If you think about registering and logging in at sites from those categories, are there any differences between the categories?
4. How do you remember the usernames and passwords for the websites you visit?
5. All of you probably know or have heard about Facebook. Facebook offers the possibility to use you Facebook account to log in at other websites. Have you previously used this functionality? Why or why not? Are there particular websites for which you would more readily use this possibility?
6. If you could have the most perfect log-in method (not registration method) you can imagine, how would that look like? *Hints: too many passwords to remember? password manager (with master-secret)? Password theft (e. g. Playstation Network, meetOne).*
7. If we look at these ideas we just discussed, what would be your favorite log-in procedure?

Key Questions

8. The new personal identity card has a rather comfortable option to register and log in with many different online services and websites. This "electronic identity" functionality can be used in a way that no personal data is disclosed

about yourselves and only an anonymous identification token is transmitted. That means that the a service you registered with then only knows during the log-in that you are the same person you were when you registered, but nothing else. No data can be collected about you and the registration is entirely anonymous. Additional data, such as your name, date of birth, email address, and so on, can be included at your discretion as it is currently possible when you register with for example a blog or another website. Hence, if you could use every online service with your new identity card, you would never need to remember usernames and passwords again. You would only need to have your identity card with you and remember the one PIN for the card.

Now, would you use such a functionality? What would cause you to use or not use this functionality?

9. For which categories of services or individual services could you imagine to use the electronic identity functionality? For which services is this functionality completely unsuited?
10. I will give each of you cards with logos and brief descriptions of eleven websites. Would you register and log-in at each of these sites with your identity card if you already have the new one or if you were to get one? Please sort the cards in three categories: Websites at which you would use the electronic identity functionality, those at which you would not, and those for which you are unsure. *[Wait]* Which proprties of the websites influenced your decisions?
11. Have you heard about the new identity card in the media? If so, which? Did those reports influence your decisions with the cards?
12. Which possibilities would you see to convince people to use their new electronic identity card for secure registration and log-in on the Internet? What would have to change?

Ending

13. Considering all the things we just talked about, how would a perfect online registration and log-in mechanism look like now?
14. I'll briefly summarize our discussion: *[Present quick summary of discussion]* Does this adequately represent your views?
15. The goal of this study was to find out why username and password work so well for authenticating with websites and why other options, such as the new personal identity card, are not adopted at all. If you think back to our discussion, did we miss any aspects? Is there anything you would have liked to talk about that we didn't?
16. Is there anything you want to add?

Thank you very much for your participation.

B Appendix: Message Encryption Study

B.1 Pre-Test Questionnaire Items

Since when have you been using Facebook?
Choose one answer: For 1 month, For 6 months, For 1 year, For 2 years, Longer, I don't know, n/a.

How often have you forgetten your Facebook password in the last 12 months?
Choose one answer: Never, Once, Twice, Three times, More than three times, n/a, I don't know, Other.

How important is it to you that only you and the recipient can read private messages?
Rate from 1 (unimportant) to 5 (important).

How often do you normally use Facebook on average?
Choose one answer: Less than an hour per day, 1 to 2 hours per day, 2 to 4 hours per day, More than 4 hours per day, More than once per week, Once per week, Monthly, Less frequently than once per month, n/a.

Approximately how many friends do you have on Facebook?

How many Facebook messages do you send per week on average?

How many of these messages have more than one recipient?

How many of these messages do you consider worthy of protection?

How often do you use the chat on Facebook?
Choose one answer: More than once per day, On a daily basis, On a weekly basis, Less than once per week, Never, n/a.

How easy do you think it is for the following persons or organisations to read your private messages on Facebook?

Rate from 1 (very easy) to 5 (very hard) for the following: Friends, Hackers, Facebook employees, Advertising Companies, US government, German government.

How high do you think the motivation is for the following persons or organisations to read your private messages on Facebook?
Rate from 1 (very low) to 5 (very high) for the following: Friends, Hackers, Facebook employees, Advertising Companies, US government, German government.

How much would it concern you if the following persons or organisations were able to read your private messages on Facebook?
Rate from 1 (very little) to 5 (very much) for the following: Friends, Hackers, Facebook employees, Advertising Companies, US government, German government.

How well do you feel you and your privacy are protected when communicating through Facebook messages?
Choose from 1 (not at all) to 5 (very well).

How well do you feel you and your privacy are protected when communicating through Facebook chat?
Choose from 1 (not at all) to 5 (very well).

B.2 Post-Task Questionnaire Items

Please rate the following questions regarding the mechanism you just used.
Choose from 1 (strongly disagree) to 5 (strongly agree) for the following:

1. I think that I would like to use this system frequently;
2. I found the system unnecessarily complex;
3. I thought the system was easy to use;
4. I think I would need the support of a technical person to be able to use the system;
5. I found the various functions in this system well integrated;
6. I thought this system was too inconsistent;
7. I would imagine that most people could learn to use this system very quickly;
8. I found the system very cumbersome to use; I felt very confident using the system;
9. I needed to learn a lot of things before I could get going with this system.

Please rate the following questions regarding the mechanism you just used.
Choose from 1 (strongly disagree) to 5 (strongly agree) for the following:

1. I would send private messages using this mechanism in the future;
2. I would send all my messages using this mechanism in the future;
3. I feel that my messages are now well protected.

B.3 Final Questionnaire Items

Please enter your age.

Please specify your gender.

Please enter your major subject.

A password is needed to use an encryption mechanism. If losing or forgetting the password led to the loss of all previous private messages, would you use such an encryption mechanism?
Choose yes or no.

Please rate the following statements with regard to the previous question about password recovery.
Choose from 1 (strongly agree) to 5 (strongly disagree) for the following: I am worried about forgetting my password; I am worried about the potential loss of all my previous messages.

Would you prefer a mechanism that is able to recover your password like it is possible on the Facebook website?
Choose yes or no.

Do you use software to encrypt your data?
Choose one or more answers: Yes, for Facebook; Yes, for email; Yes, for my hard disk; I don't know; No; Yes, for: ...

When friends have computer problems, they often ask me for help.
Choose from 1 (strongly disagree) to 5 (strongly agree).

When I have computer problems, I often ask my friends for help.
Choose from 1 (strongly disagree) to 5 (strongly agree).

What is AES?
Choose one or more answers: A browser extension; A Facebook application to store images; An encryption mechanism; I don't know; Something else: ...

Do you have any comments on this study, the procedure, the technologies used or anything else?

C Appendix: Security Measures in the Wild

C.1 Online-Survey Questionnaire

Smartphone Risk Attitudes

- **IF CODE LOCK:** Please estimate how many times you approximately unlock your phone on an average day. – *Numeric answer*
- **IF CODE LOCK:** Please briefly state why you are using a lock screen on your device. – *Open-ended answer*
- **ELSE:** Please briefly state why you chose not to use a PIN, password, or pattern lock screen on your device. – *Open-ended answer*
- **IF CODE LOCK:** Please rate the following statements concerning your lock screen. – *5-point numeric scale anchored at don't agree and fully agree.*
  - Unlocking my phone is annoying sometimes.
  - I like the idea that my phone is protected from unauthorized access.
  - It is difficult to unlock my phone.
  - I wish there was an easier way of unlocking my phone.
  - Unlocking my phone is easy.
  - I am concerned that someone might be observing my unlocking password/pattern/PIN in order to access my phone at a later time.
- What's the worst thing that could happen to your smartphone?
  - Losing the phone itself, because I would have to buy a new one.
  - Losing the data that is on my phone (e.g. photos, contacts).
  - Someone being able to access my data when I lose my phone.
  - Someone being able to abuse my accounts and apps when I lose my phone.
  - Someone being able to access my data when my phone is unattended.
  - Someone being able to abuse my accounts and apps when my phone is unattended.
  - Other: *text field*

- Please rate how the following events compare to the worst thing that could happen to your smartphone (Your answer was: <previous answer>). – *5-point numeric scale anchored at worse, similar and not as bad.*
  - Losing data on my computer
  - Losing my wallet
  - Losing the key to my home
  - Losing the key to my car
  - Getting my email account hacked
  - Someone breaking into my home
- Please rate how serious you find the following incidents. – *5-point numeric scale anchored at not serious and very serious.*
  - *same items as "What's the worst thing..."*
- How likely do you believe it is that each of the following things occurs to you personally? – *5-point numeric scale anchored at very unlikely and very likely.*
  - *same items as "What's the worst thing..."*
- How frequently do you think about each of the following things? – *5-point numeric scale anchored at very infrequently and very frequently.*
  - *same items as "What's the worst thing..."*
- How likely do you consider the following groups of people to be attempting to access your smartphone? – *5-point numeric scale anchored at very unlikely and very likely.*
  - Unknown malicious person
  - Unknown curious person
  - Known malicious person
  - Known curious person
- **IF known person considered likely:** Which of the following groups of known people did you just consider as potentially interested in accessing your phone without your permission? – *Choice from: Potentially curious person, potentially malicious person, I did not consider this group of people.*
  - Acquaintances
  - Close friends
  - Friends of friends
  - Parents
  - Children
  - Other relatives
  - Co-workers and colleagues

  - Other people

Extra Measures

- Do you sometimes take additional measures to protect your smartphone in particular situations? – *Choose all that apply.*
  - I leave my phone in a safe place before going somewhere.
  - I conceal my smartphone in my clothes or in a bag.
  - I enable a lock screen for this situation or choose a harder PIN/password/pattern.
  - Other: *text field*
- **IF MEASURES TAKEN:** Please list up to three situations in which you sometimes take additional measures to protect your smartphone. – *Open ended answer in three text fields.*
- **IF CODE LOCK:** If you think someone is able to see the screen of your phone, do you sometimes take additional measures to protect your smartphone? – *Choose all that apply.*
  - I cover my smartphone while entering my PIN or pattern.
  - I wait a moment before entering my PIN or pattern.
  - I turn around before entering my PIN or pattern.
  - I tilt my screen away before entering my PIN or pattern.
  - I change my PIN/password/pattern after someone could have seen my screen.
  - Other: *textfield*

Critical Incidents You indicated that someone had unwanted access to your smartphone. If this happened more than once, please answer this and the following questions with regard to the most severe case of unwanted access.

- Who had unwanted access to your smartphone? – *Open-ended answer in text field.*
- Please briefly described what happened during this unwanted access. – *Open-ended answer in text field.*
- Please briefly describe which harmful consequences, if any, arose from this unwanted access. – *Open-ended answer in text field.*
- What good, if any, came as a result of this unwanted access? – *Open-ended answer in text field.*
- What do you think made the unwanted access possible? – *Open-ended answer in text field.*

C.2 Online-Survey Codeplan

C.2.1 Reasons for Using Code Lock

1. Protect from specific attacker
   a) Coworker
   b) Spouse
   c) Roommate
   d) Own children
   e) Other *unwanted* individual/Stranger
   f) Unspecified people
   g) Friends
2. Protect information
   a) In general/entire phone
   b) Private/personal/sensitive information
   c) Generally *confidential* information
   d) (Confidential) *Work* info
   e) Emails/Messages
   f) Photos
   g) Contacts
   h) Calendar
   i) Other app-content
3. Protect from specific scenarios
   a) Phone protected if stolen
   b) Phone protected if lost
   c) Phone protected if misplaced
   d) Phone protected if left unattended
   e) Someone casually picking up the phone
   f) Unwanted disclosure, Pranks
   g) “Messing up” the phone
4. Protect certain action
   a) Calls
   b) Internet use
   c) Using services

   d) Play with phone
   e) Deletion
   f) Accidental input
   g) Accidental calls
   h) Other accidental use
   i) Stealing data
5. Lock is mandatory
   a) Forced by employer
   b) Forced because of custom certificate
6. Context
   a) Work
   b) Sleep
   c) Death
7. Given protection goal
   a) Increase difficulty of access
   b) Increase time to recover/find phone
   c) Access control
   d) “Safety”/Security
   e) Privacy
   f) Encrypt data
8. Other
   a) Set by default
   b) Having a lock is a habit
   c) Allows second wallpaper
   d) Previous bad experience
   e) Peace of mind
   f) Don’t know
   g) Curiosity
   h) Used to Locking
9. Off Topic/Other
10. “Protection”, Unspecific/general

C.2.2 Reasons for Not Using Code Lock

1. Inconvenience
    a) It's a hassle/annoying/easier without
    b) Mental burden
    c) Takes too much time/want instantly available
    d) Use it too frequently
    e) Don't feel like it/Just don't like it
    f) Too impatient
    g) Not eyes-free
    h) Used to existing system
2. Dislike
    a) Passwords
    b) Unlocking in general
3. No threat
    a) General: Don't need security/not concerned about security
    b) Nothing to hide/not worried about privacy
    c) No sensitive data on phone
    d) Not afraid of loosing phone
    e) Keep physically secured/never leave unattended
    f) Trust people around me/no one who wants to access
    g) Use only in private environment
    h) Phone not valuable
    i) No bad experiences so far
4. Locking may cause problems
    a) May forget my password/PIN/pattern
    b) Child may lock parent out of own phone
    c) Want finder to be able to contact me
    d) Phone accessible in emergency
    e) Shared use
5. No specific reason/Carelessness
    a) Didn't consider it/think about it
    b) Haven't gotten around to set it up yet
    c) Don't care

   d) Don't know how to set it up
   e) Don't know if available
   f) Laziness
6. Technical Reasons
   a) Phone doesn't support lock (sic)
   b) Broken Screen
   c) Slows down phone
7. Protect phone using another measure
   a) Use locking only in specific situations
   b) Rely on remote locking
   c) Leave phone at home
   d) App-specific lock
8. Rightful punishment
9. Off topic/other
10. No protection possible/is not secure anyway

C.2.3 Situations

These codes were attached to statements in which participants mentioned where they take extra measures.

1. Public spaces
   a) "Out", General public space
   b) Events (Sport, Concert)
   c) Airport
   d) Public transport (plane, train, bus)
2. Semi-Public Spaces
   a) Gym/Sports/Workout/exercise
   b) Party/Club/Bar
   c) Work/School
   d) Shopping
   e) Restaurant
   f) Cinema
3. Private spaces
   a) Home

   b) Car
4. Unknown Spaces
   a) Travel/Vacation
   b) Unfamiliar places
5. (Hardware-)Risky Conditions
   a) Water (Swimming, Boat, Rain)
   b) Sports
   c) Dirt (Beach, Cooking, Mow the lawn)
   d) Jail
   e) Lifting objects
6. Crowds
   a) General crowded places
   b) High foot-traffic area
7. Clothing
   a) No Pockets
   b) Other
8. Persons
   a) Suspicious/nosy persons
   b) Unknown/Untrusted persons
   c) Family, Kids
   d) Ex-Partner
   e) Coworkers/Other pupils
   f) General other people
   g) Friends
   h) Partner (girlfriend, boyfriend, spouse)
9. Uncontrolled Situations
   a) General less cautious situation
   b) General unattended
   c) Left charging
   d) Drinking/Socializing
   e) Sleeping
   f) Checked bags/Airport Security
10. Discomforting Environment

   a) Night/badly lit places
   b) Dangerous neighborhood/somewhere sketchy
11. Device sharing
12. Data
   a) Inappropriate
   b) Sensitive
13. Long idle times
14. Not at home
15. Activity
   a) Walking
   b) Quick errand
   c) Exercising
   d) Lodging/overnight stay
16. Off Topic/Other

C.2.4 Extra Measures

These codes were attached to statements in which participants mentioned which additional measures they take.

1. Safer mobile storage
   a) Wear close to body (e.g. in pocket)/keep out of sight
   b) Pocket in handbag/hide in purse
   c) Zippered pocket
   d) Inside pocket
   e) Backpack
   f) Have someone else carry it
   g) Keep in hand
   h) Strapped to belt/hip
2. Safer static storage
   a) At home
   b) Leave/hide in car (e.g. glove box)
   c) Locker/Drawer
   d) Leave in hotel safe
   e) Pocket instead of purse

   f) Never leave in car
   g) Other/general
3. Technical Measures
   a) Turn off
   b) Enable lock screen
   c) Have remote wiping/find my phone enabled
   d) Encrypt data
   e) Remove memory card
   f) Extra protection for specific apps
   g) Disallow access to specific apps
   h) Mute it
   i) Remove battery
   j) Have backup
   k) Use biometrics
4. Pay extra attention
   a) Check repeatedly if phone is still there/ Monitoring phone (alerts)
   b) Use it less/minimize interaction
   c) Monitoring bystanders
   d) Don't leave unattended
5. Physical measures
   a) Sturdy/special case
   b) Protect from water
   c) Leave on highest shelf (kids)
   d) Screen protector
   e) Micro-cloth
   f) Don't give to others
   g) Other physical measure
6. Data
   a) No sensitive data
   b) Different accounts
7. General/other safe place

C.3 Mini-Questionnaires

Participants were randomly presented with one of two mini-questionnaires. One concerned risks arising during unlocking and the other concerned risks to the data on the phone in general.

C.3.1 Unlocking Questionnaire

1. Who has a view on the contents of your screen right now?
   a) Unknown Person
   b) Known Person
   c) Nobody
2. IF NOT (1) NOBODY: Please rate how likely it is that someone is watching your screen right now.
   a) 5-point numeric scale (“very unlikely” to “very likely”)
3. IF NOT (1) NOBODY: Please rate how severe it would be if this person was watching your screen right now.
   a) 5-point numeric scale (“not severe at all” to “very severe”)
4. WITH CODE LOCK: Did you try to protect your code input?
   a) Yes/No
5. WITH CODE LOCK: Would you rather not have had a code lock in this situation?

   WITHOUT CODE LOCK: Would you rather have had a code lock in this situation?
   a) 5-point numeric scale (“do not agree” to “agree”)
6. In what kind of environment are you right now?
   a) Private
   b) Semi-Public
   c) Public
7. How sensitive is the data you are going to access now?
   a) 5-point numeric scale (“not sensitive at all” to “very sensitive”)

C.3.2 Data Risk Questionnaire

1. Please rate this unlock.
   a) 5-point numeric scale (“not annoying at all” to “very annoying”)

2. Did you take any additional measures to protect your phone since last using your phone?
   a) Hidden in clothes/purse
   b) Left in a safe place
   c) Other: <Text>
3. Could someone have had unwanted access to your phone since you last used it?
   a) Yes/No
4. IF YES (3): Who could have had unwanted access?
   a) Unknown Person
   b) Known Person
5. IF YES (3): How likely do you think it is that this person actually did access the device?
   a) 5-point numeric scale (“very unlikely” to “very likely”)
6. IF YES (3): How severe would the consequences of this access be, had it actually happened?
   a) 5-point numeric scale (“not severe” to “very severe”)
7. In what kind of environment has the phone been since you last used it?
   a) Private
   b) Semi-Public
   c) Public

D Appendix: Adoption Budget

D.1 Questionnaire Overview

The questionnaire contained the following questions for each scenario:

1. *What do you think is the greatest risk/the greatest danger that arises for you personally from [scenario]?*
2. *Which additional risks/dangers arising from [scenario] do you know about? You can enter up to three additional risks/dangers.*
3. *When did you last hear about these risks/dangers from others (including media, friends and family)?* Answers: *never before, a few months ago or longer, a few weeks ago, a few days ago, recently.*
4. *Which of the following statements best describes your listing of risks/dangers arising from [scenario]?* Answers:
   - *I have entered all risks/dangers that concern me.*
   - *I have entered the most important risks/dangers, but there are more.*
   - *I did not enter more risks/dangers, since I don't know about any further risks or dangers.*
   - *I did not enter more or all risks/dangers, since I feel safe on the Internet.*
   - *I did not enter more risks/dangers, because all boxes were filled.*
   - *I don't want to answer this question.*
   - *Other: [textbox]*
5. *Overall, how high do you believe the risk to your wellbeing from logging in to your social network profile to be? Please enter a number between 0 (no risk) and 100 (very high risk).*
6. *Please enter up to four consequences that may arise from the risks/dangers of [scenario] you provided in the previous question. Please begin with the most severe possible consequence and leave the additional boxes empty if you do not know any further consequences.*
7. *With regard to risks and dangers in other situations of your life, how severe do you consider your most severe consequence "[given consequence]", arising from [scenario], to be?* Answers: *(1) not severe at all to (10) very severe.*

8. *What do you think is the probability of the most severe consequence "[given consequence]" to happen to you personally?* Answers: *(1) very improbable to (10) very probable.*

D.2 Codeplan and Counts

Revised codeplan used for the final round of coding. The numbers next to the items denote the incidence of each code in the students and MTurk deployment respectively. Top level items summarize the counts of all sub-items. Note that these numbers can be higher than the sum of the contained items, as very general responses were counted towards the top-level item.

- Account Abuse – 175/128
  - Stealing credentials (unspecific) – 65/59
  - Stealing credentials (specific) – 92/60
  - Account abuse – 16/6
  - Using account for criminal purposes – 1/0
  - Endangering other accounts – 1/3
- Fraud – 106/118
  - Identity theft – 12/93
    - SSN stolen – 0/2
  - Non-existent merchandise or services – 18/5
  - Low-quality or faked merchandise or services – 8/1
  - Insufficient information on merchant – 2/1
  - Hidden costs – 17/1
- Financial Risks – 110/138
  - Theft/abuse of credit card or banking details (no account access) – 85/127
  - Abuse of online banking (mentioned phishing) – 7/1
  - Abuse of online banking (no phishing) – 8/8
  - Erroneous money transfer – 10/2
- Privacy – 220/157
  - Loss of privacy – 22/22
  - Stealing private information – 67/58
  - Leaving a trail of data – 2/7
    - Personal info stored on third-party server – 0/3
    - Need to give private info to service provider – 0/2

  - Profiling – 10/13
  - Public disclosure of private information – 12/11
  - Passing private information on to third parties – 31/15
  - Information is hard to delete online – 0/0
  - Surveillance – 36/17
    - Government Surveillance – 4/2
    - Companies – 3/0
  - Collection of data in general – 1/0
  - Abuse of personal data – 36/17
    * Abuse by other users of the same service – 1/0
    * Abuse of online photos – 0/2
- Malware and Hackers – 124/125
  - Receiving spam – 0/5
  - Malware infection – 85/61
    * "Drive by Download" – 2/5
  - Abuse of PC for illegal activities by third parties ("Botnets") – 3/3
  - Targeted attacks from unknown third parties ("Hackers") – 33/51
  - Abuse of one's IP-address – 1/0
- Psychological and Societal Risks 25/26
  - Cybermobbing, Bullying – 5/7
  - Psychological issues due to unsuitable content – 0/0
  - Internet addiction – 7/3
  - Being influenced by ads – 1/1
  - Getting depressed – 1/0
  - Account abuse to discredit someone – 0/1
  - Unpleasant social contacts – 3/3
  - Getting distracted from (more important) things – 3/6
  - Being dependent on IT services – 1/0
  - Loosing social contacts – 3/2
  - Influencing politics – 1/0
  - Loss of productivity – 3/0
- Real-world Crime – 14/47
  - Endangering one's kids – 0/1

  - Copyright violation – 1/0
  - Mixed up in a crime – 1/1
  - Stalking, Internet Predators – 8/13
  - Theft of physical things – 2/22
  - Burglary due to known absence from Internet sources – 2/9
- Health Risks – 78/102
  - Risk to health and wellbeing – 76/83
    - Meeting with serial killer – 1/6
    - Obesity – 1/0
    - Health risks because of repetitive motions and sitting – 0/6
- Own Mistakes/Negligence – 5/11
  - Insecure passwords – 1/2
  - Leaving an account logged-in – 0/4
  - Overspending – 3/5
- Misc. – 57/31
  - "General Risk" - 2/0
  - Others changing data – 2/1
  - Unreliability of other people – 44/25
  - Unreliability of services – 4/1
  - Exhausting bandwidth/data plan limit – 0/1
  - Faulty software/programming/services – 5/3
- Negative Codes – 1308/1097
  - N/A – 1115/906
  - This scenario does not apply to me – 10/3
  - Don't know – 6/0
  - Not a risk – 49/65
  - Off topic – 10/11
  - There is no risk – 18/11
  - Duplicate risk – 100/101

Bibliography

[1] Agosto, D. E., Abbas, J., and Naughton, R. Relationships and Social Rules: Teens' Social Network and Other ICT Selection Practices. *JASIST 63*, 6 (2012), 1108–1124. [p. 32]

[2] Anderson, J., Diaz, C., Bonneau, J., and Stajano, F. Privacy-enabling Social Networking over Untrusted Networks. In *Proceedings of the 2nd ACM Workshop on Online Social Networks* (2009), pp. 1–6. [p. 52]

[3] Anderson, R. J. *Security Engineering - A Guide to Building Dependable Distributed Systems (2. ed)*. Wiley, 2008. [p. 6, 7]

[4] Au, K. W. Y., Zhou, Y. F., Huang, Z., and Lie, D. PScout: Analyzing the Android Permission Specification. In *Proc. CCS* (2012). [p. 127]

[5] Baden, R., Bender, A., Spring, N., Bhattacharjee, B., and Starin, D. Persona: An Online Social Network With User-defined Privacy. In *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication* (2009), pp. 135–146. [p. 52]

[6] Beato, F., Kohlweiss, M., and Wouters, K. Scramble! Your Social Network Data. In *Proceedings of the 11th International Conference on Privacy Enhancing Technologies* (2011), Springer, pp. 211–225. [p. 52]

[7] Beautement, A., Sasse, M. A., and Wonham, M. The Compliance Budget: Managing Security Behaviour in Organisations. In *Proc. New Security Paradigms Workshop (NSPW)* (2008). [p. 21, 97, 99, 104]

[8] Benenson, Z., Gassmann, F., and Reinfelder, L. Android and iOS Users' Differences Concerning Security and Privacy. In *Proc. CHI extended abstracts* (2013). [p. 16]

[9] Bennett, C. J., and Lyon, D. *Playing the Identity Card: Surveillance, Security and Identification in Global Perspective*. Routledge, 2008. [p. 29]

[10] Besmer, A., Lipford, H. R., Shehab, M., and Cheek, G. Social Applications: Exploring A More Secure Framework. In *Proc. SOUPS* (2009). [p. 125]

[11] Bianchi, A., Oakley, I., Kostakos, V., and Kwon, D. S. The Phone Lock: Audio and Haptic Shoulder-surfing Resistant PIN Entry Methods for Mobile Devices. In *Proc. TEI* (2011), pp. 197–200. [p. 68, 69]

[12] Biddle, R., Chiasson, S., and Van Oorschot, P. Graphical passwords: Learning from the first twelve years. *ACM Comput. Surv. 44*, 4 (Sept. 2012), 19:1–19:41. [p. 67, 69]

[13] Blackwell, A. F., Church, L., and Green, T. The Abstract is 'an Enemy'. In *Proc. Psychology of Programming Interest Group (PPIG) Workshop* (2008). [p. 123, 125]

[14] Blythe, J., Camp, J., and Garg, V. Targeted Risk Communication for Computer Security. In *Proc. IUI* (2011). [p. 16, 23, 97]

[15] Blythe, J., and Camp, L. J. Implementing Mental Models. In *Proc. SPW* (2012), IEEE. [p. 25]

[16] Bødker, S. When second wave hci meets third wave challenges. In *Proc. NordiCHI* (2006). [p. 4]

[17] Bonneau, J. The Science of Guessing: Analyzing an Anonymized Corpus of 70 Million Passwords. In *Proc. IEEE S&P* (2012), pp. 538–552. [p. 26]

[18] Bonneau, J., Herley, C., van Oorschot, P. C., and Stajano, F. The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes. In *Proc. IEEE S&P* (2012), pp. 553–567. [p. 16, 26]

[19] Bravo-Lillo, C., Cranor, L., and Downs, J. Poster: What is Still Wrong With Security Warnings: A Mental Models Approach. In *Proc. SOUPS* (2010). [p. 23, 125]

[20] Bravo-Lillo, C., Cranor, L., Downs, J., and Komanduri, S. Bridging the Gap in Computer Security Warnings: A Mental Model Approach. *IEEE Security & Privacy 9*, 2 (2011), 18–26. [p. 16, 23, 125, 137]

[21] Bravo-Lillo, C., Cranor, L. F., Downs, J., Komanduri, S., Reeder, R. W., Schechter, S., and Sleeper, M. Your Attention Please: Designing Security-Decision UIs to Make Genuine Risks Harder to Ignore. In *Proc. SOUPS* (2013). [p. 16, 23, 123]

[22] Bravo-Lillo, C., Cranor, L. F., Downs, J., Komanduri, S., and Sleeper, M. Improving Computer Security Dialogs. In *Proc. INTERACT* (2011). [p. 16, 23]

[23] Brenner, M., Perl, H., and Smith, M. How Practical is Homomorphically Encrypted Program Execution? An Implementation and Performance Evaluation. In *Proc. TrustCom* (2012). [p. 13]

[24] Brooke, J. SUS: A "Quick and Dirty" Usability Scale. In *Usability Evaluation in Industry*, P. Jordan, B. Thomas, B. Weerdmeester, and A. McClelland, Eds. Taylor and Francis, 1996. [p. 59]

[25] Buchegger, S., Schiöberg, D., Vu, L.-H., and Datta, A. PeerSoN: P2P Social Networking: Early Experiences and Insights. In *Proc. ACM Workshop on Social Network Systems* (2009), pp. 46–52. [p. 52]

[26] Bundesamt für Sicherheit in der Informationstechnik. Technical Guideline TR-03127, 2011. [p. 30]

[27] Chen, P., and Popovich, P. *Correlation: Parametric and nonparametric measures.* Sage, 2002. [p. 62]

[28] Chenoweth, T., Minch, R., and Gattiker, T. Application of Protection Motivation Theory to Adoption of Protective Technologies. In *Proc. Hawaii International Conference on System Sciences (HICSS)* (2009). [p. 20]

[29] Cherdantseva, Y., and Hilton, J. A Reference Model of Information Assurance & Security. In *Proc. International Conference on Availability, Reliability and Security (ARES)* (2013). [p. 7]

[30] Cherubini, M., and Oliver, N. A Refined Experience Sampling Method to Capture Mobile User Experience . In *International Workshop of Mobile User Experience Research - Proc. CHI EA* (2009). [p. 79]

[31] Claar, C. L. *The Adoption of Computer Security: An Analysis of Home Personal Computer User Behavior Using the Health Belief Model.* PhD thesis, Management Information Systems, Utah State University, 2011. [p. 19, 20, 21]

[32] Clarke, N., and Furnell, S. Authenticating mobile phone users using keystroke analysis. *International Journal of Information Security 6*, 1 (2007), 1–14. [p. 69]

[33] Consolvo, S., Harrison, B., Smith, I., Chen, M. Y., Everitt, K., Froehlich, J., and Landay, J. A. Conducting In Situ Evaluations for and With Ubiquitous Computing Technologies. *International Journal of Human-Computer Interaction 12*, 1-2 (2007), 103–118. [p. 78]

[34] Cranor, L. A Framework for Reasoning About the Human in the Loop. *UPSEC* (2008). [p. 125]

[35] Cranor, L. F., and Garfinkel, S. *Security and Usability.* O'Reilly, 2008. [p. 2, 16]

[36] Crossler, R. E. Protection Motivation Theory: Understanding Determinants to Backing Up Personal Data. In *Proc. Hawaii International Conference on System Sciences (HICSS)* (2010). [p. 20]

[37] Cutillo, L., Molva, R., and Strufe, T. Safebook: A Privacy-preserving Online Social Network Leveraging on Real-life Trust. *IEEE Communications Magazine 47*, 12 (dec. 2009), 94 –101. [p. 52]

[38] Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science 35*, 8 (1989), 982–1003. [p. 18, 19, 98]

[39] De Luca, A., Frauendienst, B., Maurer, M.-E., Seifert, J., Hausen, D., Kammerer, N., and Hussmann, H. Does MoodyBoard Make Internet Use More Secure? In *Proc. CHI* (2011). [p. 23]

[40] De Luca, A., Hang, A., Brudy, F., Lindner, C., and Hussmann, H. Touch Me Once and I Know It's You!: Implicit Authentication Based on Touch Screen Patterns. In *Proc. CHI* (2012). [p. 69]

[41] De Luca, A., Harbach, M., von Zezschwitz, E., Maurer, M.-E., Slawik, B., Hussmann, H., and Smith, M. Now You See Me, Now You Don't – Protecting Smartphone Authentication from Shoulder Surfers. In *Proc. CHI* (2014). [p. 69, 95]

[42] De Luca, A., von Zezschwitz, E., Nguyen, N. D. H., Maurer, M.-E., Rubegni, E., Scipioni, M. P., and Langheinrich, M. Back-of-device Authentication on Smartphones. In *Proc. CHI* (2013). [p. 68, 69]

[43] De Paula, R., Ding, X., Dourish, P., Nies, K., Pillet, B., Redmiles, D., Ren, J., Rode, J., et al. Two Experiences Designing for Effective Security. In *Proc. SOUPS* (2005). [p. 125]

[44] Dey, A., and Weis, S. PseudoID: Enhancing Privacy in Federated Login. http://www.pseudoid.net, 2010. [p. 26, 29]

[45] Dhamija, R., and Dusseault, L. The Seven Flaws of Identity Management: Usability and Security Challenges. *IEEE Security & Privacy 6* (2008), 24–29. [p. 26, 29]

[46] Dhamija, R., Tygar, J. D., and Hearst, M. Why Phishing Works. In *Proc. CHI* (2006). [p. 16, 25]

[47] Dodson, B., Vo, I., Purtell, T. J., Cannon, A., and Lam, M. S. Musubi: Disintermediated Interactive Social Feeds for Mobile Devices. In *Proc. WWW* (2012), pp. 211 – 220. [p. 52]

[48] Downs, J. S., Holbrook, M. B., and Cranor, L. F. Decision Stragies and Susceptibility to Phishing. In *Proc. SOUPS* (2006). [p. 23]

[49] Downs, J. S., Holbrook, M. B., Sheng, S., and Cranor, L. F. Are Your Participants Gaming the System? Screening Mechanical Turk Workers. In *Proc. CHI* (2010). [p. 17, 70]

[50] Dutton, W., Blank, G., and Groselj, D. Cultures of the internet: The internet in britain. oxford internet survey 2013. Oxford Internet Institute, University of Oxford, 2013. [p. 98, 100]

[51] Egelman, S. My Profile Is My Password, Verify Me! The Privacy/Convenience Tradeoff of Facebook Connect. In *Proc. CHI* (2013). [p. 16, 123, 125]

[52] Egelman, S., Cranor, L. F., and Hong, J. You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. In *Proc. CHI* (2008). [p. 16, 23]

[53] Egelman, S., Oates, A., and Krishnamurthi, S. Oops, I Did it Again: Mitigating Repeated Access Control Errors on Facebook. In *Proc. CHI* (2011). [p. 50]

[54] Egelman, S., Tsai, J., Cranor, L. F., and Acquisti, A. Timing is Everything?: The Effects of Timing and Placement of Online Privacy Indicators. In *Proc. CHI* (2009). [p. 56]

[55] Fahl, S., Harbach, M., Muders, T., Baumgärtner, L., Freisleben, B., and Smith, M. Why Eve and Mallory Love Android: An Analysis of Android SSL (In)Security. In *Proc. CCS* (2012). [p. 14]

[56] Fahl, S., Harbach, M., Muders, T., and Smith, M. Helping Johnny 2.0 to Encrypt His Facebook Conversations. In *Proc. SOUPS* (2012), ACM. [p. 49, 101]

[57] Felt, A. P., Chin, E., Hanna, S., and Wagner, D. Android Permissions Demystified. In *Proc. CCS* (2011). [p. 12, 16, 126, 127]

[58] Felt, A. P., Egelman, S., Finifter, M., Akhawe, D., and Wagner, D. How to Ask For Permission . In *Proc. HotSec* (2012). [p. 12, 16]

[59] Felt, A. P., Egelman, S., and Wagner, D. I've Got 99 Problems, But Vibration Ain't One: A Survey of Smartphone Users' Concerns. In *Proc. SPSM* (2012). [p. 12, 16]

[60] Felt, A. P., Ha, E., Egelman, S., Haney, A., Chin, E., and Wagner, D. Android Permissions: User Attention, Comprehension, and Behavior. In *Proc. SOUPS* (2012). [p. 2, 12, 16, 124, 126]

[61] Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., and Combs, B. How Safe is Safe Enough? A Psychometric Study of Attitudes Towards Technological Risks and Benefits . *Policy Sciences 9*, 2 (1978), 127–152. [p. 24]

[62] Fishbein, M., and Ajzen, I. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* Addison-Wesley, Reading, MA, 1975. [p. 19]

[63] Flanagan, J. C. The Critical Incident Technique. *Psychological Bulletin 51*, 4 (1954), 327. [p. 71]

[64] Florencio, D., and Herley, C. A Large-Scale Study of Web Password Habits. In *Proc. WWW* (2007). [p. 2, 8, 10, 16, 26]

[65] Florencio, D., and Herley, C. Where do Security Policies Come From? In *Proc. SOUPS* (2010). [p. 9, 10]

[66] Friedman, B., Hurley, D., Howe, D. C., Nissenbaum, H., and Felten, E. Users' Conceptions of Risks and Harms on the Web: A Comparative Study. *Proc. CHI EA* (2002). [p. 100, 101, 121]

[67] Fromm, J., and Hoepner, P. The New German eID Card. In *Handbook of eID Security: Concepts, Practical Experiences, Technologies*, W. Fumy and M. Paeschke, Eds. Publicis, 2011, ch. 11, pp. 154–166. [p. 30]

[68] Garfinkel, S. L., and Miller, R. C. Johnny 2: A User Test of Key Continuity Management with S/MIME and Outlook Express. In *Proc. SOUPS* (2005). [p. 48, 50]

[69] Garg, V., and Camp, J. End User Perception of Online Risk under Uncertainty. In *Proc. HICSS* (2012). [p. 16, 24]

[70] Garg, V., Huber, L., Camp, L. J., and Connelly, K. Risk Communication Design for Older Adults. *Gerontechnology 11*, 2 (2012), 166. [p. 16, 23]

[71] Gaw, S., and Felten, E. W. Password Management Strategies for Online Accounts. In *Proc. SOUPS* (2006), ACM. [p. 10, 16, 26]

[72] Grote, J. H., Keizer, D., Kenzler, D., Kenzler, P., Meinel, C., Schnjakin, M., and Zoth, L. Vom Client Zur App. Tech. rep., Hasso Plattner Institute, 2010. [p. 29]

[73] Guha, S., Tang, K., and Francis, P. NOYB: Privacy in Online Social Networks. In *Proc. ACM Workshop on Online Social Networks* (2008). [p. 52]

[74] Harbach, M., Fahl, S., Rieger, M., and Smith, M. On the Acceptance of Privacy-Preserving Authentication Technology: The Curious Case of National Identity Cards. In *Proc. PETS* (2013), Springer. [p. 28]

[75] Harbach, M., Fahl, S., and Smith, M. Who's Afraid of Which Bad Wolf? A Survey of IT Security Risk Awareness. In *Proc. CSF* (2014). [p. 100]

[76] Harbach, M., Fahl, S., Yakovleva, P., and Smith, M. Sorry, I Don't Get It: An Analysis of Warning Message Texts. In *Proc USEC* (2013). [p. 16, 23]

[77] Harbach, M., Fahl, S., Yakovleva, P., and Smith, M. Sorry, I Don't Get It: An Analysis of Warning Message Texts. In *Proc USEC* (2013). [p. 123]

[78] Harbach, M., Hettig, M., Weber, S., and Smith, M. Using Personal Examples to Improve Risk Communication for Security and Privacy Decisions. In *Proc. CHI* (2014). [p. 125]

[79] Harbach, M., von Zezschwitz, E., Luca, A. D., Fichtner, A., and Smith, M. It's a Hard Lock Life: A Field Study of Smartphone (Un)Locking Behavior and Risk Perception. In *Proc. SOUPS* (2014). [p. 68]

[80] Hayashi, E., Das, S., Amini, S., Hong, J., and Oakley, I. CASA: Context-Aware Scalable Authentication. In *Proc. SOUPS* (2013). [p. 70, 86, 93, 94]

[81] Hayashi, E., and Hong, J. A Diary Study of Password Usage in Daily Life. In *Proc. CHI* (2011), ACM. [p. 10, 16, 26]

[82] Hayashi, E., Riva, O., Strauss, K., Brush, A. J. B., and Schechter, S. Goldilocks and the Two Mobile Devices: Going Beyond All-Or-Nothing Access to a Device's Applications. In *Proc. SOUPS* (2012). [p. 70, 93, 94, 141]

[83] Herath, T., Chen, R., Wang, J., Banjara, K., Wilbur, J., and Rao, H. R. Security Services as Coping Mechanisms: An Investigation Into User Intention to Adopt an Email Authentication Service. *Info Systems J.* (2012). [p. 18, 25, 28, 41, 98, 99]

[84] Herath, T., and Rao, H. R. Protection motivation and deterrence: a framework for security policy compliance in organisations. *European Journal of Information Systems 18*, 2 (2009), 106–125. [p. 21]

[85] Herley, C. So Long, And No Thanks for the Externalities: The Rational Rejection of Security Advice by Users. In *Proc. New Security Paradigms Workshop (NSPW)* (2009). [p. 22, 98]

[86] Hochbaum, G. M. *Public Participation in Medical Screening Programs; A Socio-psychological Study.* Public Health Service Publication, 1958. [p. 20]

[87] Hogarth, R. M., Portell, M., and Cuxart, A. What Risks Do People Perceive in Everyday Life? A Perspective Gained from the Experience Sampling Method (ESM). *Risk Analysis 27*, 6 (2007), 1427–1439. [p. 78, 80]

[88] Hogarth, R. M., Portell, M., Cuxart, A., and Kolev, G. I. Emotion and Reason in Everyday Risk Perception. *Journal of Behavioral Decision Making 24*, 2 (2011), 202–222. [p. 101, 102]

[89] Huang, D.-L., Rau, P.-L. P., and Salvendy, G. Perception of Information Security. *Behaviour & Information Technology 29*, 3 (2010), 221–232. [p. 24]

[90] Hühnlein, D., Petrautzki, D., Schmölz, J., Wich, T., Horsch, M., Wieland, T., Eichholz, J., Wiesmaier, A., Braun, J., Feldmann, F., Potzernheim, S., Schwenk, J., Kahlo, C., Kühne, A., and Veit,

H. On the Design and Implementation of the Open eCard App. In *Proc. Sicherheit 2012: Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)* (2012). [p. 29]

[91] Inglesant, P., Sasse, M. A., Chadwick, D., and Shi, L. L. Expressions of Expertness. In *Proc. SOUPS* (2008). [p. 125]

[92] Intille, S. S., Rondoni, J., Kukla, C., Ancona, I., and Bao, L. A Context-Aware Experience Sampling Tool. In *Proc. CHI-EA* (2003). [p. 78]

[93] Jakobsson, M., Chow, R., and Molina, J. Authentication - Are We Doing Well Enough? *IEEE Security & Privacy 10*, 1 (2012), 19–21. [p. 26]

[94] Jakobsson, M., Shi, E., Golle, P., and Chow, R. Implicit Authentication for Mobile Devices. In *Proc. HotSec* (2009). [p. 69]

[95] Johnson, M., Egelman, S., and Bellovin, S. M. Facebook and Privacy: It's Complicated. In *Proc. SOUPS* (2012). [p. 12]

[96] Jones, L. A., Antón, A. I., and Earp, J. B. Towards Understanding User Perceptions of Authentication Technologies. In *Proc. WPES* (2007). [p. 22]

[97] Karat, C.-M., Karat, J., Brodie, C., and Feng, J. Evaluating Interfaces for Privacy Policy Rule Authoring. In *Proc. CHI* (2006). [p. 125]

[98] Karlson, A. K., Brush, A. J. B., and Schechter, S. Can I Borrow Your Phone?: Understanding Concerns When Sharing Mobile Phones. In *Proc. CHI* (2009). [p. 69, 92]

[99] Kelley, P. G. Conducting Usable Privacy & Security Studies with Amazon's Mechanical Turk. In *Proc. SOUPS* (2010). [p. 17, 70]

[100] Kelley, P. G., Cranor, L. F., and Sadeh, N. Privacy as Part of the App Decision-Making Process. In *Proc. CHI* (2013). [p. 126, 129, 131, 137]

[101] Khot, R. A., Kumaraguru, P., and Srinathan, K. WYSWYE: Shoulder Surfing Defense for Recognition Based Graphical Passwords. In *Proc. OzCHI* (2012). [p. 68, 69]

[102] Kim, S.-H., Kim, J.-W., Kim, S.-Y., and Cho, H.-G. A new Shoulder-Surfing Resistant Password for Mobile Environments. In *Proc. ICUIMC 2011* (2011), ACM, p. 27. [p. 69]

[103] King, J., Lampinen, A., and Smolen, A. Privacy: Is There an App for That? In *Proc. SOUPS* (2011). [p. 50]

[104] Kittur, A., Chi, E. H., and Suh, B. Crowdsourcing User Studies With Mechanical Turk. In *Proc. CHI* (2008). [p. 17]

[105] Klasnja, P., Consolvo, S., Jung, J., Greenstein, B. M., LeGrand, L., Powledge, P., and Wetherall, D. "When I am on Wi-Fi, I am Fearless". In *Proc. CHI* (2009). [p. 101]

[106] Krueger, R. A., and Casey, M. A. *Focus Groups: A Practical Guide for Applied Research - 4th Edition.* Sage Publications, 2009. [p. 32, 33]

[107] Kujala, S., and Miron-Shatz, T. Emotions, Experiences and Usability in Real-life Mobile Phone Use. In *Proc. CHI* (2013). [p. 79]

[108] Kumaraguru, P., and Cranor, L. F. Privacy indexes: A Survey of Westin's Studies. Tech. Rep. CMU-ISRI-5-138, Carnegie Mellon University, 2005. [p. 33, 132]

[109] Kurniawan, S., Mahmud, M., and Nugroho, Y. A Study of the Use of Mobile Phones by Older Persons . In *Proc. CHI EA* (2006). [p. 32]

[110] Lambert, A. P., Bezek, S. M., and Karahalios, K. G. Waterhouse: Enabling Secure E-mail With Social Networking. In *Proc. CHI* (2009). [p. 50]

[111] LaRose, R., Rifon, N. J., and Enbody, R. Promoting Personal Responsibility For Internet Safety. *Communications of the ACM 51*, 3 (2008), 71–76. [p. 20]

[112] Lazar, J., Feng, J. H., and Hochheiser, H. *Resarch Methods in Human-Computer Interaction.* Wiley, 2010. [p. 17, 59]

[113] Liang, H., and Xue, Y. Avoidance of Information Technology Threats: A Theoretical Perspective. *MIS Quarterly 33*, 1 (2009), 71–90. [p. 18, 20]

[114] Lipford, H. R., Besmer, A., and Watson, J. Understanding Privacy Settings in Facebook with an Audience View. In *Proc. USENIX UPSec* (2008). [p. 12]

[115] Liu, Y., Gummadi, K. P., Krishnamurthy, B., and Mislove, A. Analyzing Facebook Privacy Settings: User Expectations vs. Reality. In *Proc. ACM IMC* (2011). [p. 12]

[116] Lucas, M. M., and Borisov, N. FlyByNight: Mitigating the Privacy Risks of Social Networking. In *Proc. WPES* (2008). [p. 52]

[117] Maguire, J., and Renaud, K. You Only Live Twice or "the Years We Wasted Caring About Shoulder-surfing". In *Proc. BCS Interaction Specialist Group Conference on People and Computers* (2012), British Computer Society. [p. 67, 69]

[118] Malone, D., and Maher, K. Investigating the Distribution of Password Choices. In *Proc. WWW* (2012). [p. 26]

[119] Margraf, M. The New German ID Card. In *ISSE 2010: Securing Electronic Business Processes* (2011), N. Pohlmann, H. Reimer, and W. Schneider, Eds. [p. 30]

[120] Maurer, M.-E., De Luca, A., and Hussmann, H. Data Type Based Security Alert Dialogs. In *Proc. CHI EA* (2011). [p. 23]

[121] Maurer, M.-E., De Luca, A., and Kempe, S. Using Data Type Based Security Alert Dialogs To Raise Online Security Awareness. In *Proc. SOUPS* (2011). [p. 56]

[122] Möller, A., Kranz, M., Schmid, B., Roalter, L., and Diewald, S. Investigating Self-Reporting Behavior in Long-Term Studies. In *Proc. CHI* (2013). [p. 79]

[123] Morgan, D. L. *Focus Groups as Qualitative Research.* Sage Publications, 1996. [p. 32]

[124] Motiee, S., Hawkey, K., and Beznosov, K. Do Windows Users Follow the Principle of Least Privilege? Investigating User Account Control Practices. In *Proc. SOUPS* (2010). [p. 23]

[125] National Institute of Standards and Technology (NIST). Advanced Encryption Standard (AES) (FIPS PUB 197), October 2001. [p. 51]

[126] Ng, B.-Y., Kankanhalli, A., and Xu, Y. C. Studying Users' Computer Security Behavior: A Health Belief Perspective. *Decision Support Systems 46*, 4 (2009), 815–825. [p. 20]

[127] O'Gorman, L. Comparing Passwords, Tokens, and Biometrics for User Authentication. *Proceedings of the IEEE 91*, 12 (Dec 2003), 2021–2040. [p. 69]

[128] Pahnila, S., Siponen, M., and Mahmood, A. Employees' Behavior towards IS Security Policy Compliance. In *Proc. Hawaii International Conference on System Sciences (HICSS)* (2007). [p. 21]

[129] Pandita, R., Xiao, X., Yang, W., Enck, W., and Xie, T. WHYPER: Towards Automating Risk Assessment of Mobile Applications. In *Proc. USENIX Security Symposium* (2013). [p. 126]

[130] Perito, D., Castelluccia, C., Kaafar, M., and Manils, P. How Unique and Traceable Are Usernames? In *Proc. PETS* (2011). [p. 26, 29]

[131] Poller, A., Waldmann, U., and Vowé, S. Electronic Identity Cards for User Authentication – Promise and Practice. *IEEE Security & Privacy 10*, 1 (2012), 46–54. [p. 31, 40]

[132] Rader, E., Wash, R., and Brooks, B. Stories as Informal Lessons About Security. In *Proc. SOUPS* (2012). [p. 123]

[133] Raja, F., Hawkey, K., Hsu, S., Wang, K., and Beznosov, K. A Brick Wall, a Locked Door, and a Bandit: A Physical Security Metaphor for Firewall Warnings. In *Proc. SOUPS* (2011). [p. 16, 23, 125]

[134] Raja, F., Hawkey, K., Jaferian, P., and Beznosov, K. It's Too Complicated, So I Turned It Off! Expectations, Perceptions, and Misconceptions of Personal Firewalls . In *Proc. SafeConfig* (2010). [p. 23]

[135] Riva, O., Qin, C., Strauss, K., and Lymberopoulos, D. Progressive Authentication: Deciding When to Authenticate on Mobile Phones. In *Proc. USENIX Security* (2012). [p. 69]

[136] Rogaway, P., and Wagner, D. Comments to NIST concerning AES Modes of Operations: CTR-Mode Encryption. National Institute of Standards and Technologies, 2000. [p. 51]

[137] Rogers, R. W. A Protection Motivation Theory of Fear Appeals and Attitude Change. *The Journal of Psychology 91*, 1 (1975), 93–114. [p. 19]

[138] Rosenbaum, S., Cockton, G., Coyne, K., Muller, M., and Rauch, T. Focus Groups in HCI: Wealth of Information or Waste of Resources? In *Proc. CHI* (2002), ACM. [p. 32]

[139] Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., and Tomlinson, B. Who Are the Crowdworkers?: Shifting Demographics in Mechanical Turk. In *Proc. CHI EA* (2010). [p. 17]

[140] Sasse, M. A., Brostoff, S., and Weirich, D. Transforming the 'Weakest Link' – A Human/Computer Interaction Approach to Usable and Effective Security. *BT Technology Journal 19*, 3 (2001), 122–131. [p. 2, 21, 97]

[141] Schaeffer, N. C. Asking Questions About Threatening Topics: A Selective Overview. In *The Science of Self-report: Implications for Research and Practice*, A. A. Stone, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, and V. S. Cain, Eds. Psychology Press, 1999, pp. 105–121. [p. 102]

[142] Schaub, F., Deyhle, R., and Weber, M. Password Entry Usability And Shoulder Surfing Susceptibility on Different Smartphone Platforms. In *Proc. MUM* (2012). [p. 67, 69]

[143] Schaub, F., Walch, M., Könings, B., and Weber, M. Exploring the Design Space of Graphical Passwords on Smartphones. In *Proc. SOUPS* (2013). [p. 16]

[144] Schechter, S., Dhamija, R., Ozment, A., and Fischer, I. The Emperor's New Security Indicators. In *Proc. IEEE S&P* (2007). [p. 14, 23]

[145] Shahzad, M., Liu, A. X., and Samuel, A. Secure unlocking of mobile touch screen devices by simple gestures: You can see it but you can not do it. In *Proc. MobiCom* (2013), pp. 39–50. [p. 69]

[146] Sheng, S., Koranda, C., Hyland, J., and Broderick, L. Poster: Why Johnny Still Can't Encrypt: Evaluating the Usability of Email Encryption Software. In *Proc. SOUPS* (2006). [p. 48, 50]

[147] Sun, S.-T., Pospisil, E., Muslukhov, I., Dindar, N., Hawkey, K., and Beznosov, K. What Makes Users Refuse Web Single Sign-On? An Empirical Investigation of OpenID. In *Proc. SOUPS* (2011), ACM. [p. 29]

[148] Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., and Cranor, L. F. Crying Wolf: An Empirical Study of SSL Warning Effectiveness. In *Proc. USENIX Security* (2009). [p. 2, 16, 23]

[149] Tamviruzzaman, M., Ahamed, S. I., Hasan, C. S., and O'brien, C. epet: When cellular phone learns to recognize its owner. In *Proc. SafeConfig Workshop* (2009), pp. 13–18. [p. 69]

[150] Tari, F., Ozok, A. A., and Holden, S. H. A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In *Proc. SOUPS* (2006), pp. 56–66. [p. 69]

[151] Telekom/T-Systems, D. Sicherheitsreport 2013. `http://www.telekom.com/medien/konzern/198366`, Aug 2013. [p. 98, 100]

[152] The Diaspora Project. `http://diasporafoundation.org/` – last access: 27.10.11, 2011. [p. 52]

[153] Tversky, A., and Kahneman, D. Availability: A Heuristic For Judging Frequency and Probability. *Cognitive Psychology* (1973). [p. 113]

[154] von Zezschwitz, E., Dunphy, P., and De Luca, A. Patterns in the wild: A field study of the usability of pattern and pin-based authentication on mobile devices. In *Proc. MobileHCI* (2013), pp. 261–270. [p. 69]

[155] Wang, N., Xu, H., and Grossklags, J. Third-party Apps on Facebook: Privacy and the Illusion of Control. In *Proc. Symposium on Computer Human Interaction for Management of Information Technology* (2011). [p. 50]

[156] Wash, R. Folk Models of Home Computer Security. In *Proc. SOUPS* (2010), ACM. [p. 25, 117]

[157] Wästlund, E., Angulo, J., and Fischer-Hübner, S. Evoking Comprehensive Mental Models of Anonymous Credentials. In *Open Problems in Network Security* (2012), J. Camenisch and D. Kesdogan, Eds., vol. 7039 of *LNCS*, Springer, pp. 1–14. [p. 29, 44]

[158] Weir, C. S., Douglas, G., Carruthers, M., and Jack, M. User Perceptions of Security, Convenience and Usability for Ebanking Authentication Tokens. *Computers & Security 28*, 1-2 (2009), 47–62. [p. 22]

[159] Weirich, D., and Sasse, M. A. Pretty Good Persuasion: A First Step Towards Effective Password Security in the Real World . In *Proc. NSPW* (2001). [p. 20]

[160] Whitten, A., and Tygar, J. Usability of Security: A Case Study. Tech. Rep. CMU-CS-98-155, Carnegie Mellon University, 1998. [p. 16]

[161] Whitten, A., and Tygar, J. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In *Proc. USENIX Security* (1999). [p. 2, 16, 48, 50]

[162] Wogalter, M., Dejoy, D. M., and Laughery, K. R. A Consolidated Communication-Human Information Processing (C-HIP) Model. *Warnings and Risk Communication* (1999), 15–23. [p. 125]

[163] Wogalter, M. S., Racicot, B. M., Kalsher, M. J., and Noel Simpson, S. Personalization of Warning Signs: The Role of Perceived Relevance on Behavioral Compliance. *International Journal of Industrial Ergonomics 14*, 3 (1994). [p. 125]

[164] Workman, M., Bommer, W. H., and Straub, D. Security lapses and the omission of information security measures: A threat control model and empirical test. *Computers in Human Behavior 24*, 6 (2008), 2799–2816. [p. 20]

[165] Wu, M., Miller, R. C., and Garfinkel, S. L. Do Security Toolbars Actually Prevent Phishing Attacks? In *Proc. CHI* (2006), ACM, pp. 601–610. [p. 16, 23]

[166] Yazji, S., Chen, X., Dick, R. P., and Scheuermann, P. Implicit User Reauthentication for Mobile Devices. In *Ubiquitous Intelligence and Computing* (2009), D. Zhang, M. Portmann, A.-H. Tan, and J. Indulska, Eds., vol. 5585 of *Lecture Notes in Computer Science*, Springer, pp. 325–339. [p. 69]

[167] Zakaria, N. H., Griffiths, D., Brostoff, S., and Yan, J. Shoulder Surfing Defence for Recall-based Graphical Passwords. In *Proc. SOUPS* (New York, NY, USA, 2011), SOUPS '11, ACM, pp. 6:1–6:12. [p. 69]

Curriculum Vitae

Personal Information

Name	Marian Harbach
Date of birth	27.04.1985
City of birth	Bad Schwalbach, Germany

Education

since 09/2010 — **PhD Computer Science**
Leibniz University Hannover, Germany. Distributed Computing and Security Group. Areas of research: Human Factors in IT Security and Privacy, Security Technology Adoption, Risk Communication and Awareness, Android Security.

10/2004 – 08/2010 — **Diplom Informatik**
University of Marburg, Germany. Topic of thesis: "*Semantic Validation of BPEL Fragment Compositions*", published as full paper at IEEE ICSC 2010. Project work: "*Streaming Intrusion Detection System for Grid Computing Environments*", published as full paper at IEEE HPCC 2009.

02/2008 – 10/2008 — **Master of Information Technology**
Monash University, Melbourne, Australia. Specialization in Service Oriented Architectures, full scholarship from the German Academic Exchange Service. Project work: "*Open Mobile Miner: A Toolkit for Mobile Data Stream Mining*", demo at ACM SIGKDD 2009. Dean's Achievement Award and Academic Medal for Excellence in Graduate and Postgraduate Coursework Study.

08/2001 – 06/2004 — **Abitur (A-levels)**
Berufliches Gymnasium Peter-Paul Cahensly Schule, Limburg/Lahn, Germany (high school with professional specialization). Specialization in Computer Science.

Experience

Workshops and Conference Program Committee Membership
Replication Track Chair, Usable Security Workshop 2014 co-located with NDSS. PC Member at GI RiskKom Workshop 2014; Reviewer for CHI Work-in-Progress in 2013 and Papers/Notes in 2014; Reviewer for Mobile HCI 2014 and NordiCHI 2014; Subreviewer for SOUPS 2014; Reviewer for Financial Crypto '13; PC member at IEEE DEST 2012, PDP 2012, and CALS Workshop 2011.

Teaching
Leibniz University Hannover, Germany. Lecturing on IT security and Usable Security and Privacy as well as Statistics and Empirical Methods; Seminars on Security and Privacy in Society (interdisciplinary course in conjunction with the School of Sociology and the School of Law) and introductory seminars on IT Security; Supervising students in Security and Usable Security project work; Advising students on Bachelor and Master theses.

Programming
Student Research Assistant at University of Marburg, Distributed Systems Group, Germany: Grid Development Tools (GDT) project (Eclipse plugin development); SecureGSM project: investigation of the possibilities of secure communication across GSM voice channels. Part-time programmer at ICO Innovative Computers, Diez/Lahn, Germany: development of a game-server rental system and a web shop solution for one of Europe's largest contact lens online retailer. Internship at Viessmann Werke GmbH & Co KG, Allendorf/Eder, Germany: Aspect-oriented programming, Java Enterprise Edition programming, IBM WebSphere Portal applications.