Human-centered Information Security and Privacy: Investigating How and Why Social and Emotional Factors Affect the Protection of Information Assets

---

Am Fachbereich Rechts- und Wirtschaftswissenschaften
der Technischen Universität Darmstadt

genehmigte

Dissertation

vorgelegt von

Anjuli Franz, M.Sc.

zur Erlangung des akademischen Grades
Doctor rerum politicarum (Dr. rer. pol.)

Erstgutachter: Prof. Dr. Alexander Benlian

Zweitgutachter: Prof. Dr. Dr. Christian Reuter

Darmstadt 2023

Anjuli Franz: *Human-centered Information Security and Privacy: Investigating How and Why Social and Emotional Factors Affect the Protection of Information Assets*

Darmstadt, Technische Universität Darmstadt.

Jahr der Veröffentlichung der Dissertation auf TUprints: 2023

Tag der Einreichung: 25.01.2023

Tag der mündlichen Prüfung: 29.06.2023

Acknowledgements

The completion of this thesis would not have been possible without the support and guidance of others. First and foremost, I would like to thank my supervisor, **Prof. Dr. Alexander Benlian**, for giving me the opportunity to earn my doctorate at his chair and for supporting me throughout my research journey in the field of IS. I am particularly grateful for the extraordinary freedom to pursue those research topics I was most fascinated by, while always receiving swift and valuable feedback that essentially paved the way for completing this thesis. Thank you very much for your continuous trust in me and my work.

I would also like to thank my co-supervisor, **Prof. Dr. Dr. Christian Reuter**, for his expertise and enthusiasm in the field of digital security and privacy, which greatly enriched our joint research project.

Furthermore, I want to express my gratitude to **Prof. Dr. Jason Thatcher**, who invited me to spend an exceptional research stay at Temple University, and helped me broaden my understanding of what constitutes an interesting research question and how to navigate academic life.

It's the people you work with that make the job fun – I would like to thank my **colleagues at the department of Information Systems and E-Services** as well as my **co-authors** and **students** for many inspiring discussions and countless memorable events that not only sharpened my research, but – perhaps more importantly – brought great joy and excitement to my daily work.

Finally, I would like to say a heartfelt thank you to my **parents**, who planted the seed of scientific curiosity from a very young age, and to my **grandparents**, who never ceased to try and understand what my "job with computers" is really about. Thank you to my **partner**, who is an excellent conversation partner for all things IS and beyond, and who spoiled me with Ottolenghi's cuisine in the weeks I spent finishing this thesis, as well as to my **friends** whom I was fortunate to make at home and around the globe. You enable me to learn, to grow, to think, and to reflect on my choices within and well beyond my Ph.D. work. I will be forever grateful for this extraordinary care, patience, and love.

Darmstadt, January 2023
Anjuli Franz

Abstract

Information systems (IS) are becoming increasingly integrated into the fabric of our everyday lives, for example, through cloud-based collaboration platforms, smart wearables, and social media. As a result, nearly every aspect of personal, social, and professional life relies on the constant exchange of information between users and online service providers. However, as users and organizations entrust more and more of their personal and sensitive information to IS, the challenges of ensuring information security and privacy become increasingly pressing, particularly given the rise of cybercrime and microtargeting capabilities. While the protection of information assets is a shared responsibility between technology providers, legislation, organizations, and individuals, previous research has emphasized the pivotal role of the user as the last line of defense. Whereas prior works on human-centered information security and privacy have primarily studied the human aspect from a cognitive perspective, it is important to acknowledge that security and privacy phenomena are deeply embedded within users' social, emotional, and technological environment. Therefore, individual decision-making and organizational phenomena related to security and privacy need to be examined through a socio-emotional lens. As such, this thesis sets out to investigate how and why socio-emotional factors influence information security and privacy, while simultaneously providing a deeper understanding of how these insights can be utilized to design effective security and privacy-enhancing tools and interventions. This thesis includes five studies that have been published in peer-reviewed IS outlets.

The first strand of this thesis investigates individual decision-making related to information security and privacy. Daily information disclosure decisions, such as providing login credentials to a phishing website or giving apps access to one's address book, crucially affect information security and privacy. In an effort to support users in their decision-making, research and practice have begun to develop tools and interventions that promote secure and privacy-aware behavior. However, our knowledge on the design and effectiveness of such tools and interventions is scattered across a diverse research landscape. Therefore, the first study of this thesis (article A) sets out to systematize this knowledge. Through a literature review, the study presents a taxonomy of user-oriented information security interventions and highlights crucial shortcomings of current approaches, such as a lack of tools and interventions that provide users with long-term guidance and an imbalance regarding cyber attack vectors. Importantly, the

study confirms that prior works in this field tend to limit their scope to a cognitive processing perspective, neglecting the influence of social and emotional factors.

The second study (article B) examines how users make decisions on disclosing their peers' personal information, a phenomenon referred to as privacy interdependence. Previous research has shown that users tend to have a limited understanding of the social ramifications of their decisions to share information, that is, the impact of their disclosure decisions on others' privacy. The study is based on a theoretical framework that suggests that for a user, recognizing and respecting others' privacy rights is heavily influenced by the perceived salience of others within their own socio-technical environment. The study introduces an intervention aimed at increasing the salience of others' personal data during the decision-making process, resulting in a significant decrease of interdependent privacy infringements. These findings indicate that current interfaces do not allow users to make informed decisions about their peers' privacy – a problem that is highly relevant for policymakers and regulators.

Shifting the focus towards an organizational context of individual security decision-making, the third study (article C) investigates employees' underlying motives for reporting cyber threats. With the aim to maximize employees' adoption of reporting tools, the study examines the effect of two tool design features on users' utilitarian and hedonic motivation to report information security incidents. The findings suggest that reporting tools that elicit a sense of warm glow, that is, a boost of self-esteem and personal satisfaction after performing an altruistic act, result in higher tool adoption compared to those that address solely users' utilitarian motivation. This unlocks a new perspective on organizational information security as a whole and showcases new ways in which organizations can engage users in promoting information security.

The second strand of this thesis focuses on the context of organizational information security. Beyond individual decision-making, organizations face the challenge of maintaining an information security culture, including, for example, employees' awareness of security risks, top management commitment, and interdepartmental collaboration with regard to security issues. The fourth study (article D) presents a measurement instrument to assess employees' security awareness. Complementary to the predominant method of self-reported surveys, the study introduces an index based on employees' susceptibility to simulated social engineering attacks. As such, it presents a novel way to measure security awareness that closes the intention-behavior gap and enables information security officers to nonintrusively monitor human vulnerabilities in real-time. Furthermore, the findings indicate that security education, training

and awareness (SETA) programs not only increase employees' awareness of information security risks, but also improve their actual security behavior.

Finally, the fifth study (article E) investigates the influence of external socio-emotional disruption on information security culture. Against the backdrop of the COVID-19 pandemic, the longitudinal study reveals novel inhibitors and facilitators of information security culture that emerged in the face of global socially and emotionally disruptive change over the course of 2020. Specifically, the study demonstrates that such disruptive events can influence information security culture negatively, or – counterintuitively – positively, depending on prerequisites such as digital maturity and economic stability.

Overall, this thesis highlights the importance of considering socio-emotional factors in protecting information assets by providing a more comprehensive understanding of why and how such factors affect human behavior related to information security and privacy. By doing so, this thesis answers calls for research that urge scholars to consider security and privacy issues in a larger social and emotional context. The studies in this thesis contribute to IS research on information security and privacy by (1) uncovering social and emotional motives as hitherto largely neglected drivers of users decision-making, (2) demonstrating how tools and interventions can leverage these motives to improve users' protection of information assets, and (3) revealing the importance of external socio-emotional factors as a thus far under-investigated influence on organizational information security. In practice, this thesis offers actionable recommendations for designers building tools and interventions to support decision-making with regard to information security and privacy. Likewise, it provides important insights to information security officers on how to build a strong and resilient information security culture, and guides policymakers in accounting for socially embedded privacy phenomena.

Zusammenfassung

Informationssysteme (IS) sind zunehmend in unser tägliches Leben integriert, z.B. durch cloudbasierte Kollaborationsplattformen, Smart Wearables oder soziale Medien. Dies hat zur Folge, dass nahezu jeder Aspekt des persönlichen, sozialen und beruflichen Lebens auf dem ständigen Austausch von Informationen zwischen Nutzenden und Anbietern von Online-Diensten beruht. Dadurch, dass Nutzende und Organisationen diesen Informationssystemen immer mehr persönliche und sensible Informationen anvertrauen, steigen jedoch die Anforderungen an die Gewährleistung von Informationssicherheit, Datenschutz und Privatheit. Der Schutz von Informationsbeständen liegt dabei in der gemeinsamen Verantwortung von Technologieanbietern, Gesetzgebung, Organisationen und denjenigen Personen, die das Informationssystem aktiv nutzen. Die bisherige Forschung hat insbesondere die zentrale Rolle der Nutzenden hervorgehoben, die die "letzte Verteidigungslinie" bilden. Hierbei wurde sich in früheren Forschungsarbeiten zur nutzerzentrierten Informationssicherheit und Privatheit hauptsächlich auf die kognitive Perspektive beschränkt. Darüber hinaus ist es jedoch wichtig, Phänomene der Informationssicherheit und Privatheit in den sozialen, emotionalen und technologischen Kontext der Nutzenden eingebettet zu betrachten. Die vorliegende Dissertation erweitert die bestehende Forschung zu nutzerzentrierter Informationssicherheit und Privatheit um die sozio-emotionale Perspektive. Fünf Studien beleuchten, wie und warum sozio-emotionale Faktoren Informationssicherheit und Privatheit beeinflussen und wie diese Erkenntnisse genutzt werden können, um wirksame Instrumente zur Unterstützung von Nutzenden bei Entscheidungen, die die Informationssicherheit und Privatheit betreffen, zu entwickeln.

Der erste Teil dieser Arbeit untersucht die individuelle Entscheidungsfindung in Bezug auf Informationssicherheit, Privatheit und Datenschutz. Tägliche Nutzerentscheidungen über die Preisgabe von Informationen, wie z. B. einer Phishing-Website sensible Anmeldedaten zu übermitteln oder einer App Zugriff auf das eigene Adressbuch zu genehmigen, haben entscheidende Auswirkungen auf die Sicherheit und Privatheit von Informationsbeständen. Forschung und Praxis haben daher begonnen, Tools und Interventionen zu entwickeln, die sicheres und privatheitbewusstes Verhalten fördern. Unser bisheriges Wissen über die Ausgestaltung und Wirksamkeit solcher Tools und Interventionen ist jedoch über eine vielfältige Forschungslandschaft verteilt. In der ersten Studie dieser Dissertation (Artikel A) wird daher, basierend auf einer Literaturrecherche, dieses Wissen systematisiert. Die Studie

stellt eine Taxonomie nutzerorientierter Interventionen zur Steigerung der Informationssicherheit vor und arbeitet entscheidende Defizite der derzeitigen Ansätze heraus, wie z. B. das Fehlen von Tools und Interventionen, die Nutzenden eine langfristige Orientierungshilfe bieten, und ein Ungleichgewicht in Bezug auf die Angriffsvektoren von Cyberangriffen. Darüber hinaus zeigt die Studie auf, dass frühere Arbeiten in diesem Bereich sich vorüberwiegend auf eine kognitive Verarbeitungsperspektive beschränken und den Einfluss von sozialen und emotionalen Faktoren vernachlässigen.

In der zweiten Studie (Artikel B) wird untersucht, wie Nutzende Entscheidungen über die Weitergabe persönlicher Informationen von Dritten treffen. Dieses Phänomen wird als Privatheitsinterdependenz bezeichnet. Die bisherige Literatur zeigt, dass Nutzende ein begrenztes Verständnis für die sozialen Auswirkungen ihrer Datenschutzentscheidungen haben. Studie B basiert auf einem theoretischen Modell, welches besagt, dass das Erkennen und Respektieren der Privatheitsrechte anderer für Nutzende stark von der wahrgenommenen Präsenz der anderen im eigenen soziotechnischen Umfeld abhängt. Sie stellt eine Intervention vor, die darauf abzielt, die Präsenz der persönlichen Daten anderer während des Entscheidungsprozesses zu erhöhen, was zu einem signifikanten Rückgang von interdependenten Privatheitsverletzungen führt. Die Ergebnisse zeigen implizit, dass derzeitige Benutzeroberflächen es Nutzenden nicht ermöglichen, fundierte Entscheidungen über die persönlichen Daten ihrer Mitmenschen zu treffen. Dies zeigt einen akuten Handlungsbedarf für Politik und Regulatorik auf.

Die dritte Studie (Artikel C) verlagert den Fokus auf individuelle Informationssicherheits-entscheidungen im organisatorischen Kontext. Sie untersucht die Beweggründe von Mitarbeitenden zur Meldung von Cyber-Bedrohungen. Mit dem Ziel, die Akzeptanz von Melde-Tools (z.B. im E-Mail-Programm integrierte Add-ons, die das Melden von verdächtigen E-Mails ermöglichen) auf Mitarbeitendenseite zu maximieren, untersucht die Studie die Auswirkung von zwei Tool-Eigenschaften auf die utilitaristische und hedonistische Motivation der Nutzenden, Sicherheitsvorfälle unternehmensintern zu melden. Die Ergebnisse deuten darauf hin, dass Melde-Tools, die ein Gefühl des sogenannten “warm glow”, d.h. eine Steigerung des Selbstwertgefühls und der persönlichen Zufriedenheit nach der Ausführung einer altruistischen Handlung, hervorrufen, zu einer höheren Akzeptanz des Tools führen als solche, die ausschließlich die utilitaristische Motivation der Nutzer ansprechen. Dies eröffnet einen neuen Blickwinkel auf organisatorische Informationssicherheit als Ganzes und zeigt neue

Wege auf, wie Organisationen Nutzende effektiv in die den Schutz der Informationssicherheit einbinden können.

Der zweite Teil dieser Arbeit befasst sich mit Informationssicherheit als organisatorisches Phänomen. Neben der individuellen Entscheidungsfindung stehen Unternehmen vor der Herausforderung, eine Informationssicherheitskultur zu etablieren, die beispielsweise das Sicherheitsbewusstsein der Mitarbeitenden, das Engagement der obersten Führungsebene und die abteilungsübergreifende Zusammenarbeit im Bereich Informationssicherheit einschließt. In der vierten Studie (Artikel D) wird ein Messinstrument zur Bewertung des Sicherheitsbewusstseins von Mitarbeitenden vorgestellt. Ergänzend zu der vorherrschenden Methode von Selbstbeurteilungsfragebögen führt die Studie einen Index ein, der auf der Anfälligkeit der Belegschaft für simulierte Social-Engineering-Angriffe basiert. Dieses neuartige Verfahren schließt die Lücke zwischen Absicht und tatsächlichem Verhalten der Mitarbeitenden und ermöglicht es Informationssicherheitsbeauftragten, menschliche Schwachstellen im Unternehmen in Echtzeit zu überwachen. Darüber hinaus deuten die Ergebnisse darauf hin, dass Trainings und Programme zur Steigerung der Informationssicherheit nicht nur das Sicherheitsbewusstsein der Mitarbeitenden, sondern auch ihr tatsächliches Sicherheitsverhalten verbessern.

In der fünften Studie (Artikel E) wird abschließend der Einfluss sozio-emotionaler Disruption auf die Informationssicherheitskultur untersucht. Vor dem Hintergrund der COVID-19-Pandemie zeigt die longitudinale Studie neuartige hemmende und unterstützende Einflussfaktoren auf die organisatorische Informationssicherheitskultur auf, die angesichts des globalen sozialen und emotionalen Umbruchs im Laufe des Jahres 2020 entstanden sind. Insbesondere zeigt die Studie, dass solche disruptiven Ereignisse die Informationssicherheitskultur eines Unternehmens negativ oder – kontraintuitiv – positiv beeinflussen können, je nach dem ob bestimmte Voraussetzungen, wie beispielsweise die digitale Reife und wirtschaftliche Stabilität des Unternehmens, erfüllt sind.

Insgesamt zeigt diese Dissertation die Bedeutung sozio-emotionaler Faktoren beim Schutz von Informationsbeständen auf, indem sie ein umfassenderes Verständnis dafür vermittelt, warum und wie solche Faktoren menschliches Verhalten im Zusammenhang mit Informationssicherheit, Datenschutz und Privatheit beeinflussen. Die Studien in dieser Arbeit leisten einen Beitrag zur IS-Forschung, indem sie (1) soziale und emotionale Motive als bisher weitgehend vernachlässigte Triebkräfte für die Entscheidungsfindung von Nutzenden aufdecken, (2) aufzeigen, wie Tools und Interventionen diese Motive einsetzen können, um

Nutzende beim Schützen von Informationsbeständen effektiv zu unterstützten, und (3) die Bedeutung externer sozio-emotionaler Faktoren als einen bisher wenig untersuchten Einfluss auf die Informationssicherheit in Organisationen aufzeigen. In der Praxis bietet diese Arbeit umsetzbare Empfehlungen für Designer, die Tools und Interventionen zur Unterstützung der Entscheidungsfindung in Bezug auf Informationssicherheit, Datenschutz und Privatheit entwickeln. Darüber hinaus liefert sie wichtige Erkenntnisse für Informationssicherheits-beauftragte, wie eine starke und widerstandsfähige Informationssicherheitskultur aufgebaut werden kann, und gibt politischen Entscheidungstragenden Einblicke in sozial bedingte Risiken hinsichtlich Datenschutz und Privatheit.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CISO Chief information security officer

COVID-19 Coronavirus disease 2019

DDoS Distributed denial of service

DOI Decision to disclose others' information

ESI Employee security index

EU European union

ICT Information communication and technology

IoT Internet of things

IPN Interdependent privacy salience nudge

IS Information system

IT Information technology

OSINT Open source intelligence

RC Recognition of others' ownership

RE Realization of the data transfer

RQ Research question

RS Respect for others' rights

SETA Security education, training and awareness

SOR Stimulus-organism-response

SSL Secure sockets layer

TLS Transport layer security

Chapter 1: Introduction

1.1 Motivation

Over the past decades, the evolvement of the internet has transformed the nexus between information technology (IT) and its users. Users no longer simply consume information or use software to accomplish specific tasks, but have become an interconnected socio-technical system with IT, where they actively create, curate, and disseminate information through technology (Sarker et al., 2019; Lee et al., 2015). These closely intertwined "information systems" (IS) have become integral to all aspects of social, professional, and civic life (Turel et al., 2020). As the primary mediator of human-to-human connection, IS enable users to benefit from increased access to people, goods, services, and information (Conboy, 2019).

Contemporary information systems, such as cloud-based collaboration platforms, smart wearables, and social media, heavily rely on the exchange of digitized information between users and organizations (Benlian et al., 2020; Benlian et al., 2018). This poses new challenges to information security (that is, the protection of the confidentiality, integrity and availability of the information (Samonas and Coss, 2014)) and privacy (that is, users' ability to control information about themselves and to decide to what extent information is communicated to others (Westin, 1970)). For example, users, organizations and governments frequently fall victim to cyber criminals who maliciously steal, disclose or encrypt confidential data through cyber attacks such as phishing, ransomware, or exploiting security vulnerabilities in devices or networks (Verizon, 2022). In 2022 alone, cyber criminals created over 4 billion unique phishing websites and distributed them via fraudulent emails or social media messages, attempting to trick users into disclosing sensitive information such as passwords or credit card details (APWG, 2022). Such attacks can have serious consequences for individuals, including fraud, blackmail, and loss of data, and can significantly impact organizations (e.g., through reputational damage, industrial espionage, or production shutdown) and societies (e.g., through unavailability of critical infrastructure or political upheaval) (Wright and Marett, 2010; Desouza et al., 2020; Willison and Warkentin, 2013). Novel IS landscapes, such as the Internet of Things (IoT), multiply the risk that comes with successful cyber attacks (Abomhara and Køien, 2015). Among German CEOs, 59% have stated that they see cyber attacks as the biggest business risk (PwC, 2022).

From a privacy perspective, using online services that require the disclosure of sensitive information, such as location or health data, carries inherent information privacy risks (Lowry

et al., 2017). Once shared with platforms or other individuals online, users have limited control over the storage, processing and dissemination of their own personal information (Linsner et al., 2022). This was demonstrated in the Vastaamo data breach, in which the confidential treatment records of thousands of Finnish psychotherapy patients were stolen by hackers, leading to extortion of patients by the perpetrators (The Guardian, 2020). Furthermore, whereas collecting and processing personal data can provide businesses with new sources of economic value through increasingly refined targeting options (Benlian et al., 2022; Esteve, 2017), it can also pose a threat to democracy when used to influence users' opinion, as seen during the 2016 US elections and the Brexit campaign (Isaak and Hanna, 2018; Bennett, 2016), and can lead to discrimination based on race, gender, or income (Spiekermann et al., 2022).

Against this backdrop, information security and privacy lie at the center of IS research (Lowry et al., 2017). While information security and privacy rely on a comprehensive chain of systems, including technologies, legislation, processes, and people (Lowry et al., 2017), prior literature has emphasized the pivotal role of the user, who, at the center of the target, builds the last line of defense (e.g., Reuter et al., 2022; Wang et al., 2017; Zimmermann and Renaud, 2019). In this thesis, I examine the human dimension of security and privacy through two highly relevant lenses: (1) individual information decision-making related to information security and privacy, and (2) organizational information security. Subsequently, I will outline the four research questions that guide the structure of this thesis.

1.2 Research Questions

In order to design effective measures to protect against information security and privacy risks, it is essential to understand how users perceive and react to them (Kirlappos and Sasse, 2011). The first part of my thesis therefore focuses on individual decision-making related to information security and privacy. Previous literature in this field has primarily adopted two perspectives. Both perspectives examine users' underlying decision-making process through a cognitive processing lens, either by viewing users as rational actors who carefully weigh the costs and benefits of their decisions (e.g., Schuetz et al., 2020; Moody et al., 2018; Johnston et al., 2016) or by exploring the cognitive shortcuts than can lead to biased and error-prone decision-making (e.g., Dennis and Minas, 2018; Dinev et al., 2015; Luo et al., 2013). While the cognitive processing lens has undoubtedly provided valuable insights to this field of research, the role of socio-emotional factors in shaping users' perceptions and behaviors related to information security and privacy has received limited attention in the literature. However, from the user's perspective, security and privacy decisions are deeply influenced by socio-emotional

factors, such as social norms (Kamleitner and Mitchell, 2019), emotional needs (Dinev et al., 2015), or role expectations (Liu et al., 2019). When making security and privacy decisions, users might try to respond appropriately to social cues, or rely on emotional processing to derive a judgment (Olson et al., 2007). In Table 1-1, I present three practical examples to illustrate how everyday information security and privacy decisions underlie socio-emotional motives.

Everyday experience	**Embedded information security and privacy decision**	**Potential socio-emotional influence factors**
Download and use WhatsApp	Disclose own phone number to WhatsApp, allow access to address book (i.e., disclose contacts' personal data)	• Social norm that WhatsApp is where communication happens • Perceived self-entitlement to contacts' personal data, as they are saved in one's own address book • Perceived expectation of sharing baby photos with family group chat in role as a parent
Receive an email that might be a phishing attack	Open attachment that could potentially compromise device or network	• Fear of missing out on important information • Curiosity about whether anti-virus program actually works
Report a suspicious phone call to the IT department	Proactively protect organizational information security	• Seeking attention or validation • Anxiety of bothering IT department with a minor problem • Feeling exhausted due to exceptional situations, such as the start of a global pandemic

Table 1-1. Exemplary socio-emotional factors in security and privacy decision-making

Considering such socio-emotional factors becomes particularly relevant as new technologies, such as the metaverse and smart home assistants, become more closely integrated with users' everyday socio-emotional experiences (Oh et al., 2023; Pradhan et al., 2019). To better understand the influence of socio-emotional factors on users' decision-making related to information security and privacy, my first research question aims to address this gap:

RQ1: How do socio-emotional motives affect individuals' decision-making with regard to information security and privacy?

Gaining an understanding of how socio-emotional factors drive security and privacy behavior offers promising opportunity to extend our knowledge on tools and interventions designed to increase individual security and privacy. By tools, I refer to technological means that

individuals can use to proactively improve information security and privacy at any given moment (e.g., email reporting tools, password managers). By interventions, I refer to methods that are introduced to the user at a specific moment with the goal of influencing security and privacy behavior, for example, trainings or additions to the choice architecture, such as warning about potentially dangerous content, nudging towards privacy hygiene, or offering action possibilities that promote security-aware behavior (e.g., Volkamer et al., 2017; Silic and Lowry, 2020; Schuetz et al., 2020; Almuhimedi et al., 2015). Again, previous tool and intervention approaches by research and practice have primarily addressed the user as a rational, cognitively controlled actor. For example, current anti-phishing interventions largely build on our understanding of cognitive processing to assist individuals to dynamically allocate attention while evaluating electronic messages or URLs (e.g., Petelka et al., 2019; Jensen et al., 2017b; Vance et al., 2018). However, many of these tools and interventions have been criticized for being time-consuming, effortful, and ultimately ineffective in triggering behavior that increases security and privacy (Kirlappos and Sasse, 2011; Sasse, 2015). Based on the insights from *RQ1*, this thesis seeks to investigate how our understanding of socio-emotional influence factors can be utilized to design more effective, user-oriented tools and interventions:

RQ2: How can information security and privacy tools and interventions leverage individuals' underlying socio-emotional motives to improve their security and privacy behavior?

While understanding individual behavior is crucial to secure users' personal assets and control over their privacy, new challenges arise when we look at information security from an organizational perspective. Organizations today heavily depend on the confidentiality, integrity, and availability of their data. For instance, data breaches can cause substantial reputational damage (Imprivata, 2020) and ransomware attacks can incapacitate entire industry supply chains (Tidy, 2021). Additionally, organizational assets are a lucrative target for cyber criminals, resulting in a significant increase in cyber attacks in recent years (Verizon, 2022; ENISA, 2022). As such, the second part of my thesis focuses on human-centered organizational information security.

Organizational security research has introduced the concept of security awareness, which refers to the extent to which employees understand and adhere to their organization's information security policies, rules, and guidelines (Siponen, 2000). While employees' understanding of these policies can be assessed through self-reported surveys (Kruger and Kearney, 2006), this method involves two significant shortcomings: First, surveys require significant time and effort from employees, limiting the frequency with which they can be deployed. Second, employees

who intend to comply with security guidelines might not necessarily be able to translate these intentions into actual behavior. Therefore, previous research has called for the development of non-intrusive assessment tools that address this intention-behavior gap (Lebek et al., 2014). My thesis addresses this call by posing the following research question:

RQ3: How can we measure employees' information security awareness, considering their actual security behavior?

Beyond the employee-centric perspective of security awareness, this thesis employs the concept of information security culture as a comprehensive construct to encompass the human factor in organizational information security. Information security culture is a result of employees' assumptions, values, and beliefs towards the protection of information assets, as "directed by the vision of senior management, […], influenced through internal and external factors, [and] supported by an adequate information and communication technology (ICT) environment" (Da Veiga et al., 2020, p. 19). Prior research has investigated the impact of internal factors, such as security education, training and awareness (SETA) programs (Haeussinger and Kranz, 2017; Lebek et al., 2014; Posey et al., 2015) and organizational merger (Dhillon et al., 2016), on information security culture. However, our understanding of how and why external, socio-emotional factors impact information security culture is limited (Da Veiga et al., 2020). The outbreak of the coronavirus disease 2019 (COVID-19) in early 2020 has disrupted all of the above described key components of information security culture: management priorities have changed due to economic uncertainty (Lai and Wong, 2020; Pereira et al., 2021), staff have moved from the office to remote work and a different ICT environment (Waizenegger et al., 2020; Dickinson, 2020), and employees have faced tensions related to home-schooling, job insecurity, and health concerns (Pradies et al., 2021). Moreover, cyber criminals have taken advantage of the highly exceptional situation by targeting employees working outside normal security protections, using social engineering tactics tailored to the emotionally charged climate (Naidoo, 2020). The impact of such unprecedented social and emotional disruption on organizational security culture is not yet clear, and is thus the focus of my last research question:

RQ4: How does organizational information security culture respond to external socio-emotional disruptions, such as the COVID-19 pandemic?

To address these four research questions, I conducted a literature review and four empirical studies that were published in five peer-reviewed IS outlets. In the following, I will provide an overview of the theoretical background on information security and privacy, followed by the

positioning of this thesis and presentation of the underlying research model. Subsequently, I will present the overall thesis structure and synopsis.

1.3 Theoretical Background

Information security and privacy are not isolated phenomena within the field of IS, but rather concepts that are deeply woven into the fabric of human behavior, organizational structures, and technological advancements (Lowry et al., 2017). The following subsection dives deeper into definitions of information security and privacy. Subsequently, I give a brief overview of relevant literature on individual security and privacy behavior, as well as organizational information security.

1.3.1 Information Security and Privacy

Information security and privacy are related in that they both involve the protection of information. However, they are distinct concepts that belong to different domains (Belanger et al., 2002; Biselli and Reuter, 2021). Information security is defined as "the protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide confidentiality, integrity, and availability" (NIST, 2011, p. B-5). It is a data-centric process and primarily concerned with protecting business information assets (Von Solms and Von Solms, 2005). In contrast, information privacy is a more complex, multifaceted concept that pertains to the protection of information related to a subject's identity (Solove, 2008; Knijnenburg et al., 2022). It is person-centric and is considered more of a right than a process (Pavlou, 2011). Historically, privacy has been viewed as "the ability of individuals to control when, to what extent, and how information about the self is communicated to others" (Ellison et al., 2011, p. 20). However, privacy becomes more challenging to navigate in networked environments such as social media, where one's audience and social context frequently change and blur (Knijnenburg et al., 2022). This is why modern perspectives on privacy have started to view privacy as a process of managing interpersonal boundaries with others (Petronio et al., 2021), or as an appropriate flow of information based on contextual factors such as social norms (Nissenbaum, 2004). This adds to the socio-emotional perspective on information security and privacy taken in this thesis. While confidentiality can make security a prerequisite for privacy, both concepts are fundamentally different based on the above definitions.

Due to their complex nature, the preservation of information security and privacy depends on a comprehensive chain of systems including technologies, legislation, processes, society,

economy, and people (Lowry et al., 2017). Looking at the current information security and privacy threat landscape, we find that the efficacy of technological protection measures is limited (Wang et al., 2017). Whereas tools such as firewalls, email filters, and two-factor authentication are paramount to safeguard information assets across personal and professional domains, cyber criminals often bypass such measures by using deception or social engineering tactics to exploit the human factor in order to gain access to technical systems (Mitnick and Simon, 2003; Algarni et al., 2017). This leaves the user as the last line of defense.

Similarly, current legislation does not sufficiently protect information security and privacy. For instance, the European Union (EU) has named privacy as a fundamental human right since its foundation (EUR-Lex, 2012) and has introduced the General Data Protection Regulation (GDPR) in 2018. The GDPR regulates the processing of personal data related to citizens of the European Union. It is considered best in class and has acted as a catalyst for major transformations of privacy policies worldwide (Li et al., 2019; Linden et al., 2020). Although the introduction of the GDPR has been a landmark decision for the preservation of users' personal information, it has been criticized for significant shortcomings. For example, while privacy decisions are omnipresent in increasingly sophisticated digital environments (such as deciding about cookie settings when browsing a website), the GDPR's insufficient guidance on implementing such privacy decision leaves users with nontransparent, unusable interfaces (Degeling et al., 2019; Nouwens et al., 2020). In addition, the GDPR limits its scope to a dyadic understanding of privacy (e.g., between a company and a consumer), while leaving room for gray area with regard to privacy infringements between individuals (Kamleitner and Mitchell, 2019).

Protecting information security and privacy hence often falls on the user, who can be viewed as both a critical weakness and a vital asset in the security chain, depending on the strand of literature (Zimmermann and Renaud, 2019). One way or the other, individuals' behavior substantially influences their own and others' information security and privacy. For example, employees' perception and reporting of information security incidents plays a crucial role in detecting cyber attacks that escape technological security measures (e.g., Vielberth et al., 2021; Heartfield and Loukas, 2018), organizations tremendously depend on employees promoting a culture of information security (e.g., Da Veiga et al., 2020; Alshaikh, 2020), and social media users' privacy can be affected by how their peers protect others' personal data (e.g., Symeonidis et al., 2018; Pu and Grossklags, 2016). The following subsection dives deeper into how individuals form decisions related to information security and privacy.

1.3.2 Individual Information Security and Privacy Decision-Making

Traditionally, most information security and privacy literature has assumed a rational actor making deliberate decisions (Dennis and Minas, 2018). Predominant theories in IS security literature are, for example, protection motivation theory and deterrence theory (e.g., Schuetz et al., 2020; Abbasi et al., 2021; Johnston et al., 2016; Herath and Rao, 2009), where individuals consciously weigh the costs and benefits of their behavior. In the privacy domain, this notion is reflected in the concept of the privacy calculus, where many empirical studies follow the macromodel "Antecedents–Privacy Concerns–Outcomes" (Dinev et al., 2015). As an example, it has been investigated how circumstances such as information context or intended secondary use affect users' rational evaluation of their privacy concerns (Buckman et al., 2019; Jiang et al., 2013), and how, in turn, weighing privacy concerns against, for example, social rewards or personalization benefits affects their engagement with online services (Jiang et al., 2013; Sutanto et al., 2013). Beyond these calculus models, researchers have argued that individual security and privacy decisions are subject to bounded rationality and biases (Dennis and Minas, 2018). Numerous studies have drawn on dual processing theory, which distinguishes deliberate, rational cognition from heuristic, automatic cognition (Kahneman, 2011). Prior literature has studied how triggers such as fatigue, time constraints, or limited cognitive resources move users away from high-effort information processing (i.e., security and privacy calculus) and toward low-effort information processing that is potentially flawed and biased (e.g., Dinev et al., 2015; Luo et al., 2013; Kehr et al., 2015). Low-effort cognitive processing is exacerbated by the fact that many information systems operate outside the user's awareness, such as personalized advertisements or smart speakers, which makes security and privacy issues less salient and hence more abstract and difficult to consciously decide on (Knijnenburg et al., 2022).

After a long-held focus on purely cognitive processing, recent literature has sparked a debate on the role of socio-emotional motives in shaping perceptions and decisions related to information security and privacy (e.g., Liu et al., 2019; Dinev et al., 2015; Renaud et al., 2021). Socio-emotional motives refer to cognitive and affective processes influenced by factors that emerge from the rich social and emotional contexts in which information systems operate (Nunamaker and Briggs, 2012). Initial research in this nascent field has looked at how positive and negative emotions influence employees' security precaution taking (Burns et al., 2019a), how the perceived emotional connection with information security affects employees' security learning and performance (Kam et al., 2021), and how social rules shape interdependent privacy decisions (Bélanger and James, 2020). Due to the limited body of literature in this field, prior

research has pointed out the need for developing new measurement tools to capture socio-emotional factors in users' security and privacy behavior (Renaud et al., 2021).

Regardless of the perspective used to examine individual decision-making, in practice, users are frequently left to make fundamental security and privacy decisions on their own, such as determining what information to share, which link to click, and which security incident to report (Wang et al., 2014; Kroll and Stieglitz, 2021). Prior research has begun to explore methods that aim to assist users in making these decisions. I will dive deeper into this field of literature in chapter 2 (article A), where I present the findings of a systematic literature review.

1.3.3 The Human Factor in Organizational Information Security

Organizations must safeguard their information assets from internal and external threats (Dalal et al., 2022). The human factor is a major contributor to organizational information security breaches: in 2022, 82% of organizational breaches involved the human element, such as the use of stolen credentials, phishing, misuse, or human error (Verizon, 2022). Understanding individual decision-making can aid in this effort, however, effectively managing the human dimension in organizational information security involves addressing higher-level factors beyond individual employee behavior (Wiley et al., 2020). Information security awareness and information security culture are two concepts that aim to capture the human dimension of organizational information security at large.

The construct of security awareness encompasses an individual's understanding of the goals of information security and their ability to identify and address security risks (Siponen, 2001; Bulgurcu et al., 2010). While, in the personal context, users have intrinsic motivation to protect their own information assets, employees tend to view security as an organizational problem imposed on them, not something in which they see themselves having an integral role (Johnston et al., 2019). This is natural considering that security is a secondary task for most employees, who use their ICT infrastructure for completing their everyday jobs, not for the sake of being secure (Jenkins et al., 2021). Organizations attempt to change this conception by implementing information security policies (Cram et al., 2019) and security education, training and awareness (SETA) programs that engage employees to actively contribute to information security (e.g., Silic and Lowry, 2020; Tsohou et al., 2015). However, these efforts can be complicated by the diversity of an organization's workforce, such as varying levels of security expertise (Posey et al., 2014) and vulnerability (Pienta et al., 2020) among employees. Furthermore, the dynamic and complex nature of security threats necessitates that relying on employees' sole adherence to information security policies is not sufficient to mitigate all risks. Rather, organizational

security depends on so-called extra-role behavior, such as assisting less capable colleagues or proactively reporting suspicious activities (e.g., Chen and Li, 2019; Hsu et al., 2015).

The term information security culture is an attempt to offer a more holistic view of the human factor in organizational information security (Wiley et al., 2020; Da Veiga and Eloff, 2010; Van Niekerk and Von Solms, 2005). It is defined as "all socio-cultural measures that support technical security methods, so that information security becomes a natural aspect in the daily activity of every employee" (Schlienger and Teufel, 2003b, p. 1). Given the embedded nature of information security within an organization, information security culture spans the individual, organizational, and leadership level (Da Veiga et al., 2020; Ruighaver et al., 2007). In extant literature, information security culture research hast predominantly been driven by a practice perspective. For example, prior works have developed assessment instruments (Da Veiga et al., 2020) and have investigated the effects of organizational merger (Dhillon et al., 2016) or overall organizational culture (Wiley et al., 2020) on information security culture.

1.4 Thesis Positioning

The growing number of cyber threats and the rise of privacy-invasive microtargeting capabilities pose significant risks to civil and organizational cyber space (Wright et al., 2014a; Gal-Or et al., 2018). Despite advancements in technological and legislative measures aimed at protecting information assets, the role of humans in securing these assets remains crucial (Cram et al., 2019). In response, IS research is being called upon to investigate ways to better understand and support the human aspect of information security and privacy. To address these calls, this thesis strives to demonstrate how and why socio-emotional factors influence individual and organizational security and privacy.

Figure 1-1 presents my overall research framework and the positioning of the five research articles (A-E) in this thesis. The left side of the model focuses on individual information security and privacy behavior. Drawing on environmental psychology, I position the first three articles (A-C) of this thesis within the stimulus-organism-response (SOR) model (Mehrabian and Russell, 1974). The SOR model posits that stimuli in an individual's environment influence their cognitive and affective processes (organism), which in turn influence and alter their behavior (response) (Mehrabian and Russell, 1974). In the context of this thesis, I investigate security and privacy-enhancing tools and interventions as stimuli. I first synthesize existing knowledge on user-oriented security and privacy interventions to establish a foundation for answering *RQ2* (article A). To address *RQ1*, I then explore the underlying socio-emotional

motives behind information security and privacy decision-making in two different contexts in articles B and C. Using information disclosure and information protection as key outcome variables, I examine the impact of these motives on security and privacy behavior. Finally, I use the insights gained from *RQ1* to inform the design of security and privacy-enhancing tools and interventions, hence addressing *RQ2*.

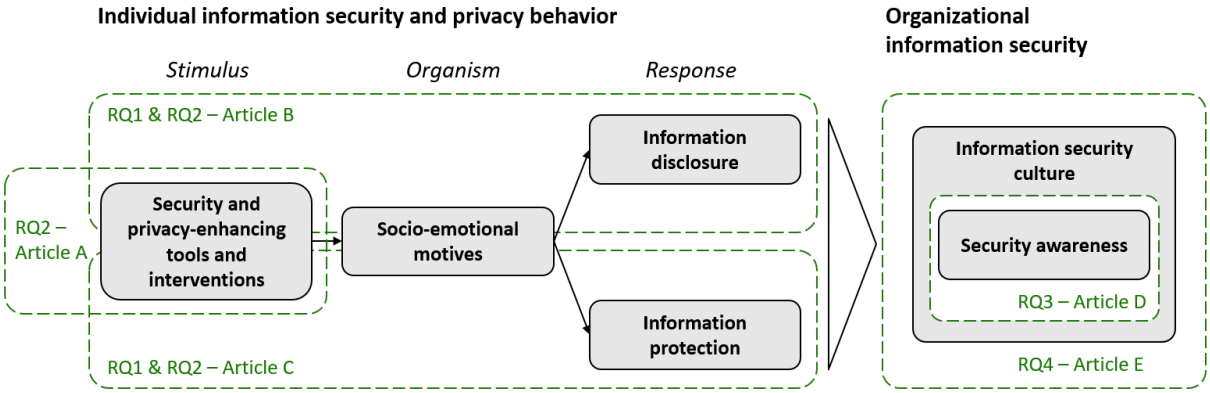


Figure 1-1. Research framework

On the right side of the model, my thesis shifts towards an organizational perspective. Following the call for a behavior-based assessment tool for employees' security awareness (Lebek et al., 2014) and addressing *RQ3*, I present an index developed in collaboration with a commercial provider of social engineering penetration tests and security awareness trainings in article D. Building on the idea that human security behavior is closely connected to users' socio-emotional context, I then examine the impact of the COVID-19 pandemic on organizational information security culture in article E, which provides valuable insights to answer my fourth research question, *RQ4*.

This thesis contributes to the literature by revealing the pivotal role of socio-emotional factors in shaping human security and privacy behavior and their implications for organizations. Additionally, the study offers practical recommendations for designers, organizations and regulators on how to promote secure and privacy-sensitive behavior among users.

1.5 Thesis Structure and Synopses

This thesis is organized into seven chapters. The introduction (chapter 1) provides background on the motivation of the thesis, introduces the overarching research questions and informs about theoretical foundations. To address the proposed research questions, five studies were conducted and published in five articles in peer-reviewed IS outlets. The five articles are each presented in chapters 2 to 6, subject to minor deviations from the originally published version

to ensure a consistent layout throughout the thesis. The concluding chapter (chapter 7) summarizes the overarching contributions to research, presents implications for practice and offers directions for future work. Table 1-2 outlines the chapters presenting the five articles of this thesis. In the following, each of the five articles is outlined by summarizing the main findings and contributions to this thesis' research questions.[1]

Chapter 2 – Article A: Taxonomy of User-Oriented Information Security and Privacy Interventions

Article A sets out to establish a foundation for the second research question, *RQ2*, through a systematic literature review. While there is an abundance of literature on interventions for supporting users' decision-making in contexts related to information security and privacy, it is fragmented across a diverse research landscape rooted in numerous disciplines such as IS, usable security, and human-computer interaction, making it difficult to gain a comprehensive understanding of different types of interventions and their effectiveness. The study focuses on the specific context of phishing as a proxy for situations where individuals are required to make decisions related to security and privacy. Phishing is a prevalent cyber threat (Verizon, 2022) that targets individuals via electronic communication with the intent of obtaining sensitive information or distributing malware (Wright et al., 2014a).

The study presents a taxonomy of user-oriented phishing interventions, which differentiates between education, training, awareness-raising and design approaches, as well as their respective subtypes. Furthermore, the article delves deeper into the interventions by examining the specific phishing attack vectors targeted, the position of the intervention within the decision-making process, and the level of user interaction required. By connecting the findings across the dimensions of analysis, the results reveal several shortcomings in prior research such as a lack of attention to attack vectors beyond fraudulent URLs, limited consideration of the intrusiveness of interventions in terms of user time and effort, and a lack of insight into the long-term effects of interventions. Interestingly, while prior research has explored different types of interventions, there is very limited research on combinations of these different types to guide users through the whole of their security and privacy decision-making journey. This study provides valuable insights on how current interventions aim to assist users in their daily information security and privacy decisions and suggests a comprehensive research agenda as a starting point for future studies.

[1] All articles use plural person pronouns (i.e., 'we'), as multiple authors were involved in their development (articles A, B, D, E) and following common practice (article C).

Chapter 2	**Taxonomy of User-Oriented Information Security and Privacy Interventions**
Article A	Franz, A.; Zimmermann, V.; Albrecht, G.; Hartwig, K.; Reuter, C., Benlian, A.; Vogt, J. (2021): "SoK: Still Plenty of Phish in the Sea – A Taxonomy of User-oriented Phishing Interventions and Avenues for Future Research", *17th Symposium on Usable Privacy and Security (SOUPS)*, August 8-13, virtual conference. **CORE[2]: B**
Chapter 3	**Underlying Mechanisms of Interdependent Privacy Decision-making**
Article B	Franz, A., Benlian, A. (2022): "Exploring Interdependent Privacy – Empirical Insights into Users' Protection of Others' Privacy on Online Platforms", *Electronic Markets*, forthcoming. **VHB[2]: B**
Chapter 4	**Utilitarian versus Hedonic Motives in Cyber Threat Reporting**
Article C	Franz, A. (2022): "Why Do Employees Report Cyber Threats? Comparing Utilitarian and Hedonic Motivations to Use Incident Reporting Tools", *43nd International Conference on Information Systems (ICIS)*, December 9-14, Copenhagen, Denmark. **VHB: A**
Chapter 5	**Behavior-based Measurement Instrument for Organizational Security Awareness**
Article D	Franz, A.; Benlian, A. (2020): "Spear Phishing 2.0: Wie automatisierte Angriffe Organisationen vor neue Herausforderungen stellen [Spear Phishing 2.0: How automated attacks present organizations with new challenges]", *HMD Praxis der Wirtschaftsinformatik*, 57(3), pp. 597-612. **VHB: D**
Chapter 6	**Information Security Culture in Times of Global Disruption**
Article E	Franz, A.; Wahl, N. (2021): "Facing Challenges Can Make You Stronger – How Global Disruptive Change Affects Organizations' Information Security Culture", *29th European Conference on Information Systems (ECIS)*, June 14-16, virtual conference. **VHB: B**

Table 1-2: Overview of the articles in this thesis

Chapter 3 – Article B: Underlying Mechanisms of Interdependent Privacy Decision-making

In article B, I address my first research question, *RQ1*, by investigating the effect of socio-emotional motives on security and privacy decisions. The study is situated within the emerging

[2] The VHB ranking was selected by the Technical University of Darmstadt as the preferred source for assessing the quality of research papers in my doctorate study program. It is published by the German Academic Association of Business Research. The VHB rankings presented in Table 1-2 are based on the version JOURQUAL3 (published in 2015), which is the latest VHB ranking at the time of writing this dissertation. Since the representation of publication outlets specific to the field of information security and privacy in the VHB ranking is limited, it was decided to additionally include the CORE2021 ranking, which provides assessments of major conferences in the computing disciplines.

field of interdependent privacy, which shifts the focus from a dyadic understanding of privacy (e.g., privacy management between a user and a company) to a perspective that acknowledges the complexity of information disclosure decisions in interconnected digital environments (Bélanger and James, 2020; Biczók and Chia, 2013). For instance, when users share group photos on social media or synchronize their address book with an online service, they implicitly make privacy decisions about their peers' personal information. However, our understanding of how such interdependent privacy decisions are formed is limited.

Article B builds on the "3R Interdependent Privacy Protection Framework" (Kamleitner and Mitchell, 2019). Kamleitner and Mitchell (2019) suggest that an interdependent privacy disclosure decision depends on three hierarchical steps: realizing the data transfer, recognizing others' ownership, and respecting others' rights. The second and third step are highly dependent on social factors. For example, users' recognition of others' ownership depends on the salience of the other in the user's socio-technical environment and can be subject to feelings of self-entitlement, and users' respect of others' rights is deeply intertwined with social norms.

To validate this framework, article B employs an online vignette experiment that simulates an interdependent privacy decision, specifically, disclosing one's address book to Instagram. The results of the experiment are analyzed using a serial mediation model, which provides empirical support for the 3R framework. Furthermore, in light of *RQ2*, the study implements an intervention aimed at increasing the salience of affected others into the decision-making process. The intervention effectively decreases the likelihood of users disclosing others' personal information by 62%. The findings confirm that users rely on socio-emotional processing to make social judgments about disclosing others' personal information. A post-hoc analysis of participants' qualitative statements provides a deeper understanding of their motives and confirms that the interdependent privacy salience intervention triggers participants' socio-emotional processing with regard to others' privacy rights. These insights are particularly relevant as current policies, such as the GDPR, do not account for interdependent privacy issues. The study hence provides important guidance for regulators and practitioners.

Chapter 4 – Article C: Utilitarian versus Hedonic Motives in Cyber Threat Reporting

Complementing the perspective of information disclosure presented in the previous article, the third article examines the proactive protection of information assets by users as a facet of their information security behavior. Rather than viewing users as the weakest link in the security chain, the study builds on the growing recognition of human users as essential contributors to information security and privacy (Zimmermann and Renaud, 2019). This is particularly relevant

in the context of cyber incident reporting, where human perception can be more effective than technological detection procedures (Heartfield and Loukas, 2018). Organizations therefore encourage their employees to report suspicious activity, such as phishing emails, to contribute to organizational information security. However, employees' usage of reporting tools is scarce.

This study again addresses both *RQ1* and *RQ2* by investigating the factors that influence employees' use of cyber incident reporting tools. Going beyond traditional perspectives of utilitarian motives, this study examines the role of hedonic drivers, specifically, the concept of "warm glow". Warm glow refers to the feeling of personal satisfaction and increased self-esteem that individuals experience after performing an apparently altruistic act (Andreoni, 1990; Iweala et al., 2019). The study employs a 2x2 online vignette experiment that offers participants an interactive reporting tool with two features designed to elicit warm glow. The results of the experiment reveal that hedonic motives are a stronger driver of employees' intention to use incident reporting tools than utilitarian motives. Additionally, the study finds that specific design choices for the reporting tool can contribute to fostering users' hedonic motivation, and hence reporting tool adoption. These findings provide a novel perspective on organizational information security at large by challenging the prevalent assumptions of why employees decide to support their organization's security efforts.

Together, articles A-C provide valuable insights into the first two research questions (*RQ1* and *RQ2*) of this thesis, focusing on how socio-emotional motives influence users in forming information security and privacy decisions, and how to leverage these motives in order to promote secure and privacy-aware behavior.

Chapter 5 – Article D: Behavior-based Measurement Instrument for Organizational Security Awareness

Article D shifts from an individual to an organizational perspective on information security and addresses the third research question *RQ3*, which is focused on measuring employees' security awareness. Prior literature largely relies on self-reported measures to assess organizational security awareness (e.g., Bulgurcu et al., 2010; Kruger and Kearney, 2006), which have been criticized for lacking realism and requiring time and effort from employees (Lebek et al., 2014). In response, this fourth article sets out to develop a non-intrusive, behavior-based measurement instrument for security awareness. In collaboration with IT-Seal GmbH, a commercial provider of social engineering penetration tests and security awareness trainings, the article introduces the "Employee Security Index" (ESI), a security awareness index based on employees' susceptibility to simulated social engineering attacks. The ESI takes into account the average time and effort invested in the attack by the cyber criminal (e.g., open source intelligence

(OSINT), registering domains, and cloning designs) and hence the knowledge and effort needed to detect the attack on the recipient side.

The article presents results from a field experiment, where the ESI was used to assess employees' security awareness in organizations and then monitor improvement after deploying SETA programs such as e-learnings. It validates prior findings from self-reported surveys in revealing that the implementation of SETA programs improves not only employees' understanding of security risks, but also actual security behavior (Haeussinger and Kranz, 2013; Chen et al., 2015b). In summary, the study showcases the capabilities of the ESI in (1) identifying unique improvement areas for specific employee groups in real time, allowing for targeted solutions, and (2) measuring the success of SETA programs. This offers guidance to information security practitioners on how to measure their employees' security awareness in a more realistic and less intrusive manner than traditional self-reported surveys.

Chapter 6 – Article E: Information Security Culture in Times of Global Disruption

Article E is crucially motivated by the emergence of the COVID-19 pandemic, which has presented not only individuals but also organizations with unprecedented challenges. The pandemic has impacted organizational information security in four important ways: (1) the sudden shift to remote work caused by the need to prevent COVID-19 infections has substantially influenced the role that ICT plays in the workplace (Carroll and Conboy, 2020; Waizenegger et al., 2020), (2) the economic challenges that many organizations have had to face has reshuffled top management's priorities and allocation of funds (Lai and Wong, 2020; Pereira et al., 2021), (3) cyber criminals have exploited the highly exceptional situation by targeting employees working outside normal security protections with social engineering tactics tailored to the emotionally charged climate (Naidoo, 2020), and (4) employees have faced tensions related to home-schooling, job insecurity, and health concerns (Pradies et al., 2021), potentially drawing attention away from security matters.

Article E addresses *RQ4* by assessing the impact of the COVID-19 pandemic on organizational information security culture. The study is based on interviews with 17 information security practitioners conducted in June and October 2020, which provide insights on how organizations' information security culture changed over the pandemic compared to pre-pandemic times. The longitudinal study allows for an examination of short- and long-term factors and reveals novel facilitators and inhibitors across the individual, organizational and leadership level. Through the lens of the punctuated equilibrium theory, the article discusses how a global disruption such as the COVID-19 pandemic can fundamentally change

information security culture, either positively or negatively. The study contributes to the organizational information security literature by extending prior theoretical frameworks and illustrates the role of digital maturity and economic stability as factors that can influence the direction of the effect of disruption on organizational information security.

Additional article (not included in this thesis):

In addition to the articles listed above, I contributed to the submission and publication of the following manuscript during my time as a Ph.D. candidate. This article, however, is not included in this thesis:

Franz, Anjuli; Croitor, Evgheni (2021): Who Bites the Hook? Investigating Employees' Susceptibility to Phishing: A Randomized Field Experiment, *29th European Conference on Information Systems (ECIS)*, June 14-16, virtual conference. **VHB: B**

Chapter 2: Taxonomy of User-Oriented Information Security and Privacy Interventions

Title:	SoK[3]: Still Plenty of Phish in the Sea – A Taxonomy of User-Oriented Phishing Interventions and Avenues for Future Research
Authors:	Anjuli Franz, Technische Universität Darmstadt, Germany Verena Zimmermann, Technische Universität Darmstadt, Germany Gregor Albrecht, Technische Universität Darmstadt, Germany Katrin Hartwig, Technische Universität Darmstadt, Germany Christian Reuter, Technische Universität Darmstadt, Germany Alexander Benlian, Technische Universität Darmstadt, Germany Joachim Vogt, Technische Universität Darmstadt, Germany
Published in:	Symposium on Usable Privacy and Security (2021), August 8-13, Virtual Conference.

Abstract: Phishing is a prevalent cyber threat, targeting individuals and organizations alike. Previous approaches on anti-phishing measures have started to recognize the role of the user, who, at the center of the target, builds the last line of defense. However, user-oriented phishing interventions are fragmented across a diverse research landscape, which has not been systematized to date. This makes it challenging to gain an overview of the various approaches taken by prior works. In this paper, we present a taxonomy of phishing interventions based on a systematic literature analysis. We shed light on the diversity of existing approaches by analyzing them with respect to the intervention type, the addressed phishing attack vector, the time at which the intervention takes place, and the required user interaction. Furthermore, we highlight shortcomings and challenges emerging from both our literature sample and prior meta-analyses, and discuss them in the light of current movements in the field of usable security. With this article, we hope to provide useful directions for future works on phishing interventions.

[3] SoK stands for "Systematization of Knowledge", which is a submission category of the Symposium on Usable Privacy and Security.

2.1 Introduction

Phishing is a frequently employed cyber attack to get hold of users' sensitive information, such as login details or banking account numbers. Furthermore, criminals increasingly use phishing attacks to distribute malware (Wright et al., 2014b). The consequences of a successful attack can reach from individual personal losses or compromised accounts to complete organizations or networks being infected by malware, often combined with ransom demands. For example, the years between 2014 and 2020 were marked by *Emotet*, a modular trojan using targeted phishing emails with weaponized Microsoft Word files (Patsakis and Chrysanthou, 2020). It is crucial to consider that phishing attacks do not primarily target hardware or software vulnerabilities, but the user – the human factor within the socio-technical system. While there are several tools and approaches that aim to identify malicious contents automatically (e.g., Tian et al., 2018; Verma and Dyer, 2015), the increasingly sophisticated and personalized nature of phishing attacks makes it hard for algorithms to detect and block phishing emails, websites, or malicious software. This leaves a large amount of responsibility to the user. However, detecting phishing attempts is not the user's first priority (Wu et al., 2006b), for instance, while using email programs: instead, users in various contexts aim to efficiently solve their tasks and answer what they perceive to be emails sent by customers or colleagues when they become victims of a phishing attack.

To enable users to be the ultimate wall of defense in cyber security, research and practice have developed a number of user-oriented interventions against phishing attacks. Among those are education and training approaches (Sheng et al., 2007; Kumaraguru et al., 2009; Canova et al., 2015), where users develop knowledge and skills that they can transfer to real-world phishing attempts. To complement these, awareness-raising measures or design considerations (Egelman et al., 2008; Marforio et al., 2016; Nicholson et al., 2017; Petelka et al., 2019) aim to guide users towards secure online behavior in situ. While developing adequate countermeasures that assist end-users in combating phishing attacks is highly relevant, finding both effective and usable user-oriented phishing interventions is still an unresolved problem (Allodi et al., 2019). Considering the diverse research landscape on phishing interventions across various research disciplines (e.g., cyber security, human-computer interaction, or social science), it is challenging to gain an overview of what types of interventions have already been investigated. The design of interventions may significantly differ between phishing attack vectors, the moment at which the intervention takes place, or approaches that increase the attention in a specific moment vs. those that encourage long-term capability to deal with phishing attacks

autonomously. To our knowledge, a comprehensive literature review of existing approaches is missing to date. We argue that a systematization of prior phishing interventions, particularly with respect to their variety across multiple characteristics, will help to identify trends and gaps in the phishing intervention literature. Furthermore, a discussion in the light of current usable security movements will lead to a better understanding of promising directions for successful user assistance in the phishing context. Our research thus aims to shed light on the following two research questions:

1: How does current research on user-oriented phishing interventions tackle the aim of guiding users towards secure online behavior?

2: Which avenues for future research emerge from the existing phishing intervention literature?

In this work, we offer a comprehensive systematization of user-oriented phishing interventions with respect to the intervention type, the addressed attack vector, the moment at which the intervention takes place, as well as the degree of user interaction. We thereby complement broader reviews such as the work of Zhang-Kennedy and Chiasson (2021), who have reviewed tools for cyber security awareness and education more generally. Our contributions are threefold: First, we present an extensive literature analysis of prior research on user-oriented phishing interventions (Brocke et al., 2009; Schryen et al., 2020), bridging the research streams of both educational and design measures. Guided by previous rudiments of phishing intervention classifications (Jansen and Schaik, 2019; Kirlappos and Sasse, 2011; Wash and Cooper, 2018; Xiong et al., 2019), we introduce a novel taxonomy of user-oriented phishing interventions consisting of four categories and ten subcategories. Second, we explore central characteristics such as the time at which the intervention takes place throughout the user's decision process, which phishing attack vectors are commonly addressed by the studied interventions, and the degree of user interaction required. Beyond that, we thirdly take into account critical considerations of leading usable security researchers (Wash, 2020; Sasse, 2015; Cranor and Garfinkel, 2005) and discuss shortcomings of prior phishing intervention approaches. In summary, we offer a novel insight into phishing intervention research and present potential avenues for future works.

2.2 Methodology

To categorize and understand the landscape of existing phishing interventions, we have conducted a systematic literature review, following the "preferred reporting items for a

systematic review and meta-analysis" (PRISMA) guideline (McInnes et al., 2018; Moher et al., 2009). Literature reviews have been argued to play an important role in developing domain knowledge, e.g., by synthesizing prior research works, identifying research gaps, and developing a research agenda (Schryen et al., 2020). To cover the diverse research landscape, our initial search comprised the databases ACM Digital Library, IEEE Xplore, and Web of Science. The search was limited to peer-reviewed studies in English that were available as of June 2020.

The search term was identical across databases and applied to the title and abstract of all included articles. For an article to be included in the analysis, it had to contain the term *phish** and one of the following terms to allow for a plurality of intervention types: *interven** OR *prevent** OR *educat** OR *detect** OR *train** OR *nudg** OR *appeal*.

In addition to the database search, we analyzed the Google Scholar top ten security conferences and journals as well as the A* and A CORE-ranked security conferences and journals. Most of them had already been included in the analyzed databases (e.g., CHI, S&P, CCS, Computers & Security). Only journals and conferences that had not been covered by the previous database search underwent an additional manual title search. These included the USENIX Network and Distributed System Security Symposium NDSS and the accompanying usable security events USEC and EuroUSEC, as well as the USENIX Security Symposium and the co-located SOUPS conference from 2014 onwards[4]. In addition to our search term-based search, we have complemented our sample with two other relevant articles that we became aware of through our literature research.

With the above-described search procedure, we have identified a total of 2,124 publications. Afterward, we have conducted a title and abstract screening to exclude irrelevant articles. Articles were excluded if they matched one of the following criteria:

- Deals with a different topic not related to phishing in the sense of cyber security
- Intervention is not user-oriented in that the user cannot see or act upon an intervention (e.g., an algorithm that invisibly filters and blocks suspicious emails)

Table 2-2 in the appendix details the distribution across the different databases before and after the title and abstract screening. After the aforementioned procedure as well as the deletion of two duplicates, a total of 80 articles remained for a detailed analysis. As for the full-text

[4] Before 2014, the SOUPS proceedings were included in the ACM database.

screening, we have read and analyzed the 80 articles independently among the authors to ensure best possible thoroughness. Since this literature review has emerged from a cross-disciplinary collaboration between seven security researchers with backgrounds in computer science, information systems and psychology, we were able to discuss the literature from various angles and finally agreed on one final review. The full-text analysis further reduced the literature count by 16 articles: First, we excluded research works that did not address a user-oriented phishing intervention in the full text (see second exclusion criterion above). Second, we excluded similar articles by the same authors (e.g., a conference paper and a subsequent, very similar journal publication), and kept only the latest and more extensive version. Our final literature sample thus includes 64 articles.

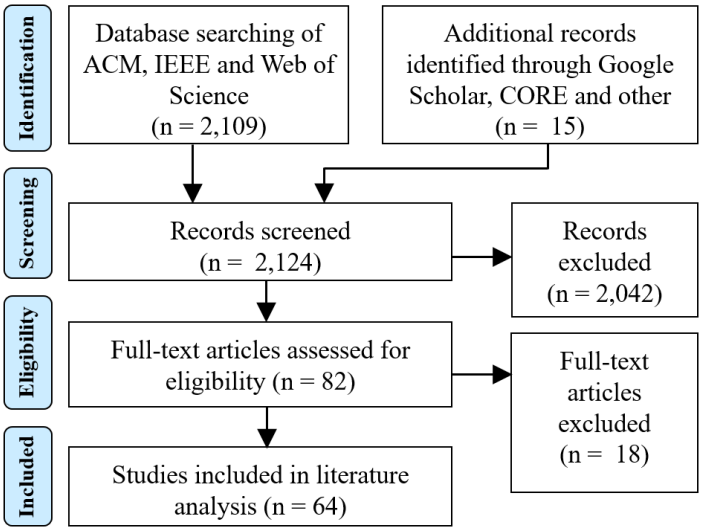


Figure 2-1. PRISMA diagram of literature screening process

Figure 2-1 shows a flowchart that details the number of screened, excluded and included articles following the PRISMA statement (McInnes et al., 2018; Moher et al., 2009).

2.3 Results

In the following, we present a detailed analysis of our literature sample. We first provide an overview of the **methodological range** employed by previous phishing intervention research (section 2.3.1). Categorizing the studied interventions with regard to their design and intended effect, we then derive a **taxonomy of user-oriented phishing interventions** (section 2.3.2). We further consider the **phishing attack vector** that the intervention aims to address (section 0), the **time at which the intervention takes place** (section 2.3.4), as well as the **degree of user interaction** (section 2.3.5).

For a comprehensive categorization of the analyzed phishing interventions across the whole literature sample, please refer to Table 2-3 in the appendix.

2.3.1 Overview of Methodological Approaches

With respect to the methodological approach, 13 research works have presented exclusively **conceptual ideas** of phishing interventions. For example, (Dhamija and Tygar, 2005) have discussed factors that make securing users against phishing a challenging design problem and have derived design requirements for authentication schemes. Studies that have gathered empirical data have drawn on **surveys** (3 publications), **lab** (20 publications), **online** (12 publications), or **field experiments** (16 publications) to analyze, e.g., the efficacy or usability of user-oriented phishing interventions. For instance, the effect of training material embedded in the process of sorting emails has been studied by (Kumaraguru et al., 2010), who have first employed a think-aloud vignette lab experiment, which has then been further tested in the field in the form of an online training game.

As for sample sizes, studies in our literature data range from small (< 20 participants) representative groups (e.g., Iacono et al., 2014; Canova et al., 2015; Wu et al., 2006b) to large-scale experiments with more than 1,000 participants (e.g., Wash and Cooper, 2018; Ronda et al., 2008; Kumaraguru et al., 2010). Field experiments were often conducted among university students and staff (e.g., Wash and Cooper, 2018), rarely among non-university employees (e.g., Reinheimer et al., 2020), or by evaluating real-world users' interactions with browser extensions or applications (e.g., Ronda et al., 2008).

While most research articles in our sample have explored short-time effects of phishing interventions, some have employed longitudinal studies in order to investigate long-term effects. For example, Kumaraguru et al. (2009) have observed knowledge retention of at least 28 days for users who had been trained via simulated phishing attacks and Silic and Lowry (2020) have employed a long-term field experiment to investigate longitudinal effects of gamification on employees' intrinsic motivation to comply with security efforts.

With regard to the validity of experimental setups, previous works have pointed out that information security behavior research heavily relies on studying users' information security behavior as their primary activity on a computer (Dennis and Minas, 2018; Hassandoust et al., 2020; Herzberg and Margulies, 2013). In reality, however, responding to phishing threats is a secondary task that is embedded in a primary task, such as answering email or searching the internet. This leads to users facing the difficulty of switching between their primary and

secondary activity, which may result in overlooking security warnings or disregarding educational offers. While many lab and online studies of our sample have studied their subjects' behavior as a primary task (e.g., by asking them to sort links into "legitimate" or "phishing" (Arachchilage and Love, 2013; Stockhardt et al., 2016)), others have assigned them fictional primary tasks to attend to. By using cover stories, such as sorting emails for a colleague or shopping online (Petelka et al., 2019; Kirlappos and Sasse, 2011), researchers have aimed to study phishing detection as a secondary task. However, it is arguable whether such artificial experimental setups can align with the complex nature of phishing. With regard to the realism of phishing experiments, Schechter et al. (2007) have shown that role-playing participants behave less securely than those who act in a personal context (e.g., participants asked to log into a bank account with predefined passwords showed less secure behavior than those using their own passwords). While online or lab experiments are essential to test and refine theories of user behavior as well as to improve artifacts in human-computer interaction, conducting studies in a realistic environment is crucial to allow for robust and practice-oriented results. In our literature sample, less than one third (16 of 51) of experiments have been conducted in a real-world field setting.

2.3.2 A Taxonomy of User-Oriented Phishing Interventions

Our literature review has revealed that, while user-oriented phishing interventions all pursue one common goal (to protect users from phishing threats), they vary widely with regard to their underlying concepts and intended effect. Prior literature has presented vague attempts of categorizations of phishing interventions. For example, Kirlappos and Sasse (2011) have described two main approaches, namely anti-phishing indicators and user education, whereas Xiong et al. (2019) have distinguished between warnings and training, and the integration of both. Similarly, Wash and Cooper (2018) has observed three styles of phishing interventions: general-purpose training messages that communicate "best practices", fake phishing campaigns, and in-the-moment warning messages. We chose to follow a fourth approach by Jansen and Schaik (2019), who have roughly described four different categories of user-oriented phishing interventions: **education**, **training**, **awareness-raising** and **design**. In their pure form, education and training interventions typically promote sustainable, long-term secure behavior, with the central aim that the application of knowledge and skills transfers to the real-world and enables users to engage in secure practices (Van Schaik et al., 2017), whereas awareness-raising and design interventions aim to improve users' security during specific activities (such as logging into a website or reading an email) in the short term. Our literature

analysis has revealed, however, that interventions often incorporate elements of more than one type.

Based on the literature data, we have derived a taxonomy of user-oriented phishing interventions as presented in Figure 2-2 and Table 2-1. In the following sections, we will describe the four categories and their respective subcategories in detail.

Education

Purely educational interventions focus on developing knowledge and understanding of phishing threats and ways to mitigate them, e.g., by providing educational media, such as texts or videos, or by discussing online threats during in-class training. For this category, we have identified 7 publications in total. However, only three of them have considered education as a solitary intervention. For example, Wash and Cooper (2018) have investigated which role the perceived origin of phishing education material plays in terms of effectiveness and have found that facts-and-advice-based training from perceived security experts surpasses the same training from peers. Four research works have studied phishing education in interaction with awareness-raising interventions by adding educational texts to fear appeals (Schuetz et al., 2020; Jansen and Schaik, 2019) or warnings (Yang et al., 2017). For example, Yang et al. (2017) have found that a warning trigger combined with an educational text enhances its effectivity, whereas the educational element itself was not sufficient to provide phishing protection. Others have first provided extensive education in order to refer back to it during awareness-raising interventions later on (Blythe et al., 2011). Education interventions have been studied in rather traditional text-based, video-based or in-class formats. More progressive formats, such as online games, comprised interactive and hands-on exercises and were hence categorized as training.

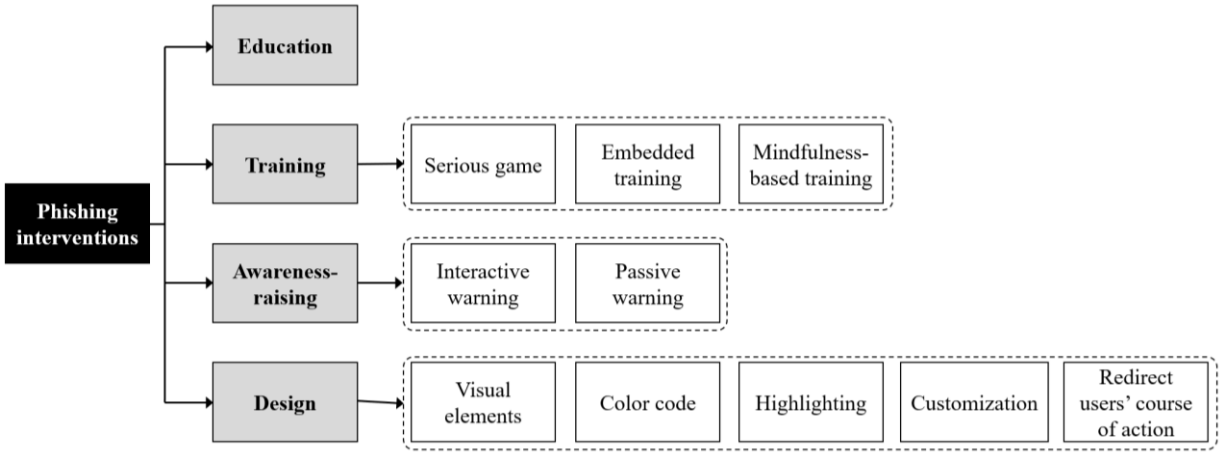


Figure 2-2. A taxonomy of user-oriented phishing interventions (overview)

Category	Definition	Phishing interventions	Articles
Education	Educational interventions aim at developing knowledge and understanding of phishing and how to protect oneself against it.		
Education		Text-based, video-based, or in-class education	(Blythe et al., 2011; Jansen and Schaik, 2019; Lastdrager et al., 2017; Reinheimer et al., 2020; Schuetz et al., 2020; Wash and Cooper, 2018; Yang et al., 2017)
Training	Training interventions refer to interactive elements or exercises, which provide users with hands-on practice. They often take place by presenting a realistic phishing attempt within a secure environment.		
Serious game	Serious games refer to gamified contexts in which users can train how to recognize and analyze phishing attacks.	Online game (e.g., "NoPhish"), mobile app, board game, escape room game	(Arachchilage and Love, 2013; Baslyman and Chiasson, 2016; Beguin et al., 2019; Canova et al., 2015; Gokul et al., 2018; Cuchta et al., 2019; Fatima et al., 2019; Hale and Gamble, 2014; Hale et al., 2015; Kumaraguru et al., 2010; Perrault, 2018; Scott et al., 2014; Sheng et al., 2007; Silic and Lowry, 2020; Weanquoi et al., 2017; Wen et al., 2019)
Embedded training	Embedded training refers to training schemes that combine testing users' behavior in their normal environment with instant corrective performance feedback.	Phishing simulation in combination with a "teachable moment" (e.g., "PhishGuru")	(Alnajim and Munro, 2009; Burns et al., 2019b; Burns et al., 2012; Caputo et al., 2013; Carella et al., 2017; Greene et al., 2018b; Kumaraguru et al., 2009; Kumaraguru et al., 2007; Kumaraguru et al., 2008; Marsden et al., 2020; Stembert et al., 2015; Wash and Cooper, 2018; Xiong et al., 2019)
Mindfulness-based training	Mindfulness-based approaches refer to trainings that increase users' awareness of context.	Approaches that teach users to dynamically allocate attention during message evaluation	(Jensen et al., 2017b)
Awareness-raising	Awareness-raising interventions refer to warnings that are placed in situ and raise users' awareness of potential phishing attempts during their primary course of action.		
Interactive warning	Interactive warnings refer to awareness-raising interventions that do require user interaction, i.e., interrupt the users' course of action.	Forced-attention warning, security questions, interactive fear appeal	(Abbasi et al., 2016; Egelman et al., 2008; Gastellier-Prevost et al., 2011; Jansen and Schaik, 2019; Petelka et al., 2019; Reeder et al., 2018; Schechter et al., 2007; Schuetz et al., 2020; Stembert et al., 2015; Volkamer et al., 2017; Wiese et al., 2018; Wu et al., 2006b; Yang et al., 2017; Yao and Shin, 2013; Yue, 2012)
Passive warning	Passive warnings refer to awareness-raising interventions that do not require user interaction.	Security toolbar, display of information on the legitimacy of a website	(Blythe et al., 2011; Wu et al., 2006a; Egelman et al., 2008)
Design	Design interventions refer to design choices that aim at supporting or guiding users' behavior with respect to their secure handling of online activities.		
Visual elements	Visual elements refer to interventions that use the	UI dressing, dynamic security skins, trust logo,	(Dhamija and Tygar, 2005; Herzberg and Jbara, 2008;

	visual appearance of, e.g., a login form or website, to support users' security behavior.	image	Herzberg and Margulies, 2013; Iacono et al., 2014; Kirlappos and Sasse, 2011; Li et al., 2012; Marforio et al., 2016; Schechter et al., 2007; Varshney et al., 2012; Yee and Sitaker, 2006)
Color code	Color codes refer to simple visual cues for users to distinguish between secure and risky environments.	Traffic light colors	(Wu et al., 2006a; Wiese et al., 2018; Kirlappos and Sasse, 2011)
Highlighting	Highlighting refers interventions that draw users' attention towards critical elements.	Domain highlighting, sender highlighting, highlighting differences in out-of-focus tabs	(De Ryck et al., 2013; Lin et al., 2011; Nicholson et al., 2017; Volkamer et al., 2017)
Customization	Customization refers to interventions that let users customize the visual appearance of, e.g., a login form.	Custom icon, custom image, custom UI dressing	(Dhamija and Tygar, 2005; Herzberg and Margulies, 2013; Herzberg and Jbara, 2008; Iacono et al., 2014; Marforio et al., 2016; Schechter et al., 2007; Varshney et al., 2012; Yee and Sitaker, 2006)
Redirect users' course of action	This category refers to interventions that redirect users' course of action, for example by offering more secure alternatives.	Browser sidebar for entering credentials, suggesting alternative websites, creating habit of using bookmarks, delayed password disclosure	(Herzberg and Margulies, 2013; Jakobsson and Myers, 2007; Miyamoto et al., 2014; Ronda et al., 2008; Wu et al., 2006b)

Table 2-1. A taxonomy of user-oriented phishing interventions (detailed)

Training

Compared with educational interventions, training goes one step further. It typically involves some kind of hands-on practice, where users develop skills that they can apply in case of a real threat. Since the term "training" is quite widespread in everyday language use, interventions that have been described as training by the respective authors might have been categorized as education within this work. Training approaches aim to enable users to identify phishing websites, phishing emails, or other malicious attacks. They employ interactive elements or exercises, where users can develop skills such as reading a URL, analyzing an email, or recognizing social engineering attempts. They often do so by exposing the user to a similar attack within a secure environment, either in an artificial or a real-world setup. Within our phishing literature data, about half of the publications (31 research papers) were dedicated to training interventions. Among them, we were able to distinguish several approaches.

Training interventions are typically rule-based. That is, their goal is to train individuals to identify certain cues to take protective action (Stockhardt et al., 2016). In our sample, 16 publications have explored such training in a **serious game** context, mostly taking place online

and often focusing on teaching users how to identify phishing links by using cues in URLs (e.g., Stockhardt et al., 2016; Sheng et al., 2007; Canova et al., 2015; Arachchilage et al., 2016). For instance, Sheng et al. (2007) have introduced "Anti-Phishing Phil", a game that is designed to teach users how to identify fraudulent websites based on the use of IP addresses, subdomains or deceptive domains in a URL. Similarly, "NoPhish" is a mobile app that guides users through several levels of analyzing and recognizing phishing URLs (Canova et al., 2015; Stockhardt et al., 2016). The authors have found a long-term effect with regard to users' knowledge retention; that is, users who had played the NoPhish game have shown a better ability to decide upon the legitimacy of a URL. Silic and Lowry (2020) have observed that gamified security training systems, which include elements such as levels or leader boards, enhance users' intrinsic motivation and yield better security behavior. Offline games have been explored in the form of board (Baslyman and Chiasson, 2016) or escape room (Beguin et al., 2019) games.

Apart from gamified contexts, **embedded training** has gained momentum in recent phishing intervention research. Embedded training describes interventions that *"train a skill using the associated operational system including software and machines that people normally use"* (Alnajim and Munro, 2009, p. 406). In other words, embedded training combines testing users' behavior in their normal personal or work environments with instant corrective performance feedback. It has been argued that the experience of "being phished" constitutes a so-called most teachable moment, where lasting change to attitudes and behaviors is possible (Caputo et al., 2013). Embedded training has been studied by 13 publications in our literature sample. As an example, "PhishGuru" is a program that simulates harmless but realistic phishing emails right into users' email inboxes (Kumaraguru et al., 2008; Kumaraguru et al., 2007; Kumaraguru et al., 2009; Kumaraguru et al., 2010). When falling for a simulated phishing attempt (i.e., clicking on a phishing link), users were redirected to a training website explaining how phishing attacks work and how they can protect themselves from fraudulent emails and websites. Embedded training is a promising approach with regard to the real-world environment it takes place in: users are not in a training environment (such as an online game), but receive training only if they fall for a simulated phishing attempt during their everyday duties. Thus, knowledge and changes in security attitudes and behaviors can be transferred to real phishing attempts more easily. This is reflected in a growing business of embedded "phishing simulation training" by commercial information security companies [5]. Kumaraguru et al. (2009) have shown that

[5] For example, Proofpoint ThreatSim (proofpoint.com), Sophos Phish Threat (sophos.com), IT-Seal Awareness Academy (it-seal.de), Lucy Security (lucysecurity.com), and many others.

training with "PhishGuru" helps users retain what they learned in the long term and that multiple training interventions increase performance.

Beyond rule-based training, Jensen et al. (2017b) have shown that expanding the rather conventional training toolkit with **mindfulness-based training** leads to a better ability to avoid phishing attacks. Mindfulness training teaches users to dynamically allocate attention during message evaluation ("*(1) Stop! (2) Think ... (3) Check.*") and aims to increase users' awareness of context. This method seems to be particularly effective for participants who were already confident in their detection ability.

Awareness-raising

The third category, awareness-raising, aims at focusing users' attention on potential threats and their countermeasures in situ, that is, as part of their primary course of action. Awareness-raising interventions might, for example, interrupt the user's workflow to set security-conscious behavior on their agenda. We have identified 17 studies of awareness-raising interventions, of which three explore **passive warnings** (i.e., the warning does not require user interaction), and 15 investigate on **interactive warnings** (i.e., the warning does require user interaction). Several prior studies have shown that passive interventions such as security toolbars in an internet browser are ineffective at preventing phishing attacks (Egelman et al., 2008; Wu et al., 2006a).

Interactive warnings have been shown to have promising effects on users' phishing vulnerability. For example, the browser sidebar "Web Wallet" (Wu et al., 2006b) acts as a secure way to submit sensitive information by suggesting alternative safe paths to intended websites and forcing users' attention by integrating security questions. Several research works have explored the mechanism of forced attention: Volkamer et al. (2017) have introduced "TORPEDO", an email client add-in that delays link activation for a short period of time. As for web browser phishing warnings, Egelman et al. (2008) have shown that interactive warnings, where users have to choose between options such as "Back to safety" or "Continue to Website", are heeded significantly more often compared to passive warnings. Furthermore, Petelka et al. (2019) have shown that link-focused warnings are more effective than general email banner warnings in protecting users from clicking on malicious URLs, and that forced attention amplifies this effect. When comparing awareness-raising interventions that include educational elements (such as descriptions of the consequences of phishing, or explanations why a certain link or file is classified as potentially dangerous) to those that do not provide any additional information, the former were found to be more effective (Stembert et al., 2015; Yang et al., 2017). Two research works have examined the potential of fear appeals, that is, short,

informative messages that communicate threats, and have found that concrete fear appeals (compared with abstract fear appeals) are more effective to increase actual compliance behavior (Schuetz et al., 2020; Jansen and Schaik, 2019). This indicates that a combination of warning, forcing users' attention, and therein embedded tangible education yields a promising protection against phishing threats.

Design

Lastly, design choices can act as phishing interventions if they facilitate desirable user behavior (Jansen and Schaik, 2019). We have identified 20 publications that investigate design interventions aimed at supporting users' secure handling of email and online activities.

Visual elements play a role in several research works (10 publications). For instance, the potential of "dynamic security skins" has been explored by Dhamija and Tygar (2005), who have presented an authentication scheme where users rely on visual hashes from a trusted source that match the website background for legitimate websites.

Visual elements also come into play when offering users design options to **customize** security indicators, such as custom images or icons. An example is "Passpet", a browser extension by Yee and Sitaker (2006) that acts as a password manager and an interactive custom indicator. Iacono et al. (2014) have proposed so-called "UI-dressing", a mechanism that relies on the idea of individually dressed web applications (e.g., by using customized images) in order to support the user in detecting fake websites.

Color codes refer to simple visual cues (e.g., traffic light colors) for users to distinguish between secure and risky environments. They have, so far, been observed to be of limited success in the form of security indicators that signal whether a website is genuine or fake (Kirlappos and Sasse, 2011; Wu et al., 2006a). Furthermore, Wiese et al. (2018) have explored color codes in the context of email application UI design, where they were used to indicate the presence of digital signatures.

In contrast, **highlighting** draws users' attention to critical elements. For example, both Volkamer et al. (2017) and Lin et al. (2011) have investigated the effectiveness of domain highlighting in order to enable users to find the relevant part of a URL, whereas Nicholson et al. (2017) have explored highlighting an email's sender name and address.

Other design interventions set out to **redirect users' course of action**, for example, by creating the habit of using browser bookmarks instead of hyperlinks to access sensitive websites such as login pages (Herzberg and Margulies, 2013). Ronda et al. (2008) have developed

"iTrustPage", a tool that warns the user about suspicious websites (e.g., a fake PayPal website). Beyond that, it offers corrective action in the form of suggesting alternative websites that are deemed trustworthy based on Google's search index (e.g., the real PayPal website).

Surprisingly, while the concept of digital nudging has gained widespread attention (among others in usable security research, e.g., Zimmermann and Renaud, 2021; Kankane et al., 2018; Coventry et al., 2014; Choe et al., 2013) in recent years, only one article in our sample has investigated the effect of a nudge: Next to highlighting the name and address of an email's sender, Nicholson et al. (2017) have investigated the effect of a social salience nudge ("62% of your colleagues received a version of this email") on users' phishing vulnerability. While several other design interventions contain nudge-like elements (such as color codes or highlighting), none of them have been designed as or labelled a nudge by the respective authors. We will further elaborate on the potential of digital nudging in phishing interventions in section 0.

2.3.3 Which phishing attack vector does the intervention address?

While the term "phishing" originally describes cyber attacks that aim for users' passwords, it is now used to describe all sorts of attack vectors (Dennis and Minas, 2018). Those attack vectors differ in terms of the criminals' intended outcome (e.g., disclosure of confidential information or implanting malware) and the user's primary action during which the attack takes place (e.g., clicking on a link or downloading a file). In the following, we will analyze the range of attack vectors that the phishing interventions in our sample aim to intervene in detail.

Phishers predominantly choose email messages as their first approach towards the user (Wash and Cooper, 2018). About 3.9 billion people worldwide have email accounts and collectively send and receive over 290 billion emails per day (The Radicati Group Inc. , 2019). Email thus presents a means of communication that can easily be abused to take advantage of users' credulity by blending into daily personal or professional correspondence. Since attackers employ social engineering techniques (e.g., urgency cues or trustworthy-seeming visual elements) to elicit specific actions such as clicking a link, opening an attachment, or disclosing sensitive information, **deceptive email messages** themselves can be considered as an attack vector. Seventeen publications address users' ability to distinguish legitimate emails from phishing emails by paying attention to the email message itself. For instance, Caputo et al. (2013) have studied embedded phishing training that aims at educating users on how to recognize phishing emails based on various criteria such as mismatched names, spelling mistakes, or intuition.

Phishing messages furthermore often offer a link, which, for example, might execute a drive-by download of ransomware (Wash and Cooper, 2018) or redirect the user to a website masquerading as a legitimate login page. Previous research suggests that, after recipients click on a phishing link, they rarely detect subsequent fraudulent attempts such as a counterfeit login page or change their course of action (Wright and Marett, 2010). **Disguised URLs** (such as, e.g., *paypa1.com*, *mybank.com-secure.biz,* or *tinyurl.com/XYZ*), that make the user believe that they are clicking on a reliable link, hence constitute a prominent attack vector. Accordingly, more than half of our literature sample (33 publications) explores user-oriented phishing interventions that aim at preventing users from clicking malicious links. These interventions mostly consider links in the context of an email. For example, Volkamer et al. (2017) email client add-on "TORPEDO" uses tooltips to focus the user's attention on a link's domain. While links with whitelisted or previously visited domains will be activated immediately when clicked, "TORPEDO" will delay the activation of other links for a few seconds to encourage the user to check the URL's domain carefully. Several training games provide users with an in-depth explanation and exercise about how URLs can be obfuscated to mimic reputable sources, and have been shown to help users make better decisions concerning the legitimacy of URLs in the long term (Stockhardt et al., 2016; Sheng et al., 2007; Canova et al., 2015).

While links are usually accessed via clicking on a link, **QR codes** gain in popularity due to their ease of distribution and fast readability. Since the user has no means to examine the URL behind a QR code before scanning it, they constitute a hidden security threat. One single publication in our sample has addressed this issue by exploring security features of QR code scanners that help users to detect phishing attacks (Yao and Shin, 2013).

Besides disguised URLs, **imitated websites** can present another attack vector. For example, cyber criminals employ imitations of well-known websites in order to exploit users' trust in visually familiar or trustworthy environments. Ten publications in our literature sample have addressed this attack vector. For example, Iacono et al. (2014) have proposed an intervention that relies on the idea that the whole appearance of a web application is dressable according to the user's individual preferences, raising users' attention for unofficial sites that do not align with the expected appearance. Regarding phishing interventions that are being displayed on websites, Kirlappos and Sasse (2011) have revealed that arbitrary logos, certifications, or advertisements that do not imply trustworthiness of a website might have a higher reassurance to users than actual security indicators. This gives an example of how user-oriented

interventions themselves can be exploited by cyber criminals to trick users into placing trust into a website.

When browsing the internet, interventions such as padlock icons or warning messages inform the user about a website's **SSL/TLS certificates**[6]. Interventions that inform or warn the user about SSL/TLS have been addressed, for example, by Reeder et al. (2018) or Schechter et al. (2007). So-called man-in-the-middle attacks, where criminals use legitimate websites that do not encrypt data transmission by SSL/TLS to capture the user's sensitive data during an online transaction, have been a serious phishing attack vector in the past. Since nowadays, however, more than 80% of phishing sites have SSL/TLS encryption enabled (APWG, 2020), this attack vector will likely cease to play a role in the near future.

We now move from the preliminary stages (such as tricking users into trusting an email, link, or website) to the centerpiece of a phishing attack. One central aspiration of cyber criminals is to lure their victim into disclosing sensitive information, e.g., login credentials. Accordingly, several prior works (12 in our literature sample) have studied interventions that address the process of users' **authentication**. For example, Dhamija and Tygar (2005) have introduced an interaction technique for authentication that provides a trusted window in the browser dedicated to username and password entry, which uses a photographic image to create a trusted path between the user and password entry fields. Similarly, Yee and Sitaker (2006)'s browser extension "Passpet" constitutes a password manager that helps users securely identify trustworthy login forms.

Besides fishing for credential data, phishers' efforts are directed at prompting the user to download or execute **malware**, that is, malicious software that can harm the user's device or their entire network. Malware attacks have rapidly grown over the recent years, e.g., in the form of ransomware attacks (Sorensen, 2018). Surprisingly, interventions that aim at preventing users from executing malware are scarce in our literature data. Only three publications have addressed this attack vector: Wen et al. (2019) have included different kinds of potentially malicious attachments in their conception of a role-play anti-phishing training game, whereas Reeder et al. (2018) have explored users' interaction with browser warnings that warn against downloading malware. Reinheimer et al. (2020) have taught how to identify dangerous files in their in-class training. Malicious **mobile applications** can act as a phishing attack vector, for

[6] Transport Layer Security (TLS), and its predecessor, Secure Sockets Layer (SSL), are cryptographic protocols designed to provide communications security over a computer network

example, by masquerading as a legitimate online banking app. One publication in our sample has discussed personalized security indicators in mobile applications (Marforio et al., 2016).

In addition to the above-described investigations of specific phishing attack vectors, 10 publications have approached the topic of phishing in a more general manner. Most of these publications have examined training formats, such as online games, that cover the phenomenon of **phishing in a broader sense** without addressing or intervening one attack vector in particular.

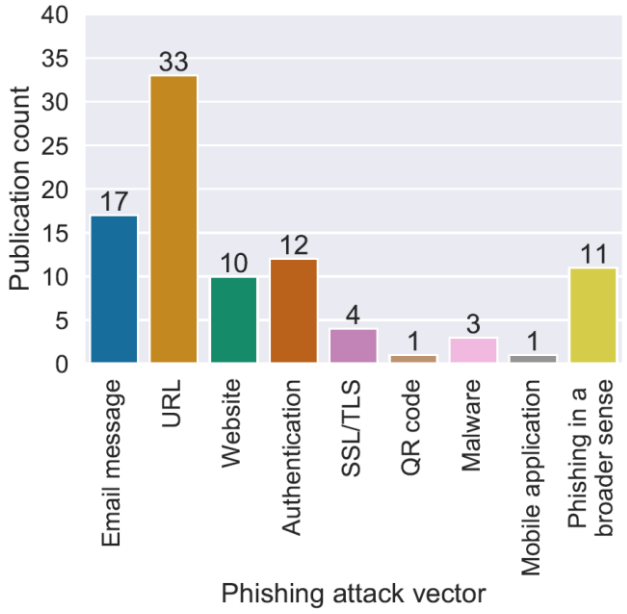


Figure 2-3. Overview of the attack vectors addressed by phishing interventions

Figure 2-3 illustrates the distribution of our literature data across different phishing attack vectors. Since some publications address interventions to more than one phishing attack vector, the sum of the displayed data points is larger than the literature sample size of 64 articles. For a detailed categorization of all articles, please refer to Table 2-3 in the appendix.

2.3.4 When Does the Intervention Take Place?

Diving deeper into the analysis of user-oriented phishing interventions, we have further considered the point in time at which the intervention takes place. We have found that many interventions, mostly those aiming at training or education, are designed to take place as a precautionary measure, often long before the user interacts with a potential phishing context. We have identified 23 articles that present such interventions and have labeled them as **pre-decision interventions**. For instance, Jansen and Schaik (2019) have shown that confronting users with fear appeal messages is suitable to heighten their cognitions, attitudes, and intentions with regard to secure online behavior. Furthermore, all kinds of non-embedded education or

training (e.g., in-class education (Lastdrager et al., 2017), online games (Sheng et al., 2007), mobile training apps (Canova et al., 2015) clearly take place pre-decision.

Most of the approaches in our literature sample focus on interventions that take place during users' course of action, that is, **during the user's decision** between phishing and legitimate content in a real-world context. Those 31 articles mostly describe awareness-raising and design interventions, sometimes combined with educational elements. For instance, Petelka et al. (2019) have examined the effectiveness of different levels of link-focused warnings when sorting emails, whereas various design interventions such as color codes, customization or highlighting aim to support users' decisions during their course of action.

We have further identified 11 publications describing interventions that take place **post-decision**, that is, after a user's decision on potential phishing contents was already made. This goes especially for embedded training, where training follows right after the user has been "phished" by a simulated attack.

Combinations of pre-, post-, and during decision intervention have been studied only once in our sample: Blythe et al. (2011) have introduced an approach that consists of initial video-based education, which is then referred back to by security warnings during the users' individual course of action.

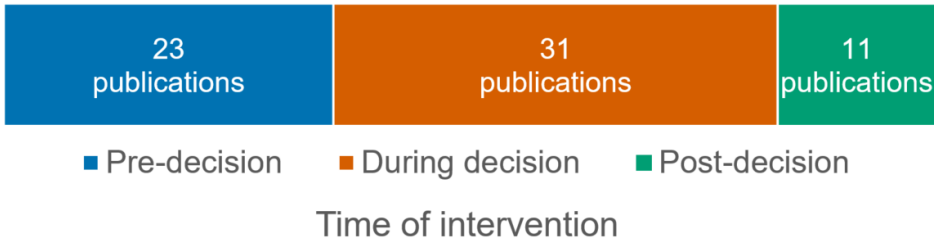


Figure 2-4. Positioning of the intervention within users' decision-making process

While several research works have employed longitudinal studies to examine the long-term effects of user-oriented phishing interventions (see section 2.3.1), little has been investigated on interventions that take place regularly, e.g., by giving regular warnings or recurrently providing users with training. Reinheimer et al. (2020) have explored the effect of reminding users of initial phishing awareness education and have found that reminders after half a year are recommended and that measures based on videos or interactive examples perform better than text-based reminders. Furthermore, several embedded training interventions have been explored in terms of the effect of recurrently simulated phishing emails (e.g., Canova et al.,

2015; Carella et al., 2017; Kumaraguru et al., 2009; Marsden et al., 2020). Figure 2-4 sums up the distribution of the time of intervention across our literature sample.

2.3.5 Does the Intervention Require User Interaction?

Beyond the categorization as presented in Table 2-1, we have analyzed all interventions in terms of whether they require active user interaction, e.g., whether the user's workflow is interrupted by the intervention and whether the user can only proceed when undertaking a certain action or decision. These interventions were classified as **interactive**. In contrast, interventions that only provide information or feedback to the user without actively interrupting their workflow are deemed **passive** interventions. Some of the 64 articles in our literature sample have addressed both interactive and passive interventions.

Across our sample, 48 publications describe phishing interventions that require user interaction. We mainly divide between two kinds of interactive interventions, one being interactive warnings, which usually require a few seconds of the user's time and attention before they can proceed with the task at hand (e.g., Petelka et al., 2019; Egelman et al., 2008). The other subset is formed by training and education approaches, which commonly require the user to actively engage in an exercise for at least several minutes up to hours, for example, online training games (Hale and Gamble, 2014; Canova et al., 2015; Sheng et al., 2007) or in-class training (e.g., Lastdrager et al., 2017). A total of 16 interventions can be described as passive, including passive warnings (e.g., Wu et al., 2006a), some educational interventions (e.g., Jansen and Schaik, 2019), and also several interventions belonging to the design category. As an example, we have classified domain highlighting (Lin et al., 2011) as passive, since it does not require any interaction on the user's side and can also be easily ignored, or even overlooked, by the user.

2.4 Discussion

In the previous section, we have examined a plethora of user-oriented phishing interventions from various angles and have revealed surprising and relevant insights. Above all, we have found a highly fragmented landscape of educational interventions, training, awareness-raising warnings, and anti-phishing designs, which users need to navigate through when being pushed towards secure online behavior. To summarize and connect the findings across the dimensions of analysis, Figure 2-5 displays an integrative plot of all phishing interventions in our sample. Since some articles have addressed several attack vectors or intervention categories, they appear more than once. Getting back to our two research questions, we devote the remainder of this

article to discussing our findings and positioning them in current usable security research. After looking at the user effort and intrusiveness of prior phishing interventions in section 2.4.1, we discuss the potential of digital nudges regarding phishing prevention in section 0. We then address the role of users' cognitive processes when dealing with potential security threats in section 2.4.3. Further, we consider the imbalance of phishing attack vectors addressed by prior intervention research in section 2.4.4, and discuss the potential of tailored phishing interventions in section 2.4.5. Subsequently, we highlight methodological aspects in section 2.4.6, and lastly address limitations of our work in section 2.4.7. We then sum up our contributions in section 2.5, including an overview of our propositions for future phishing intervention research.

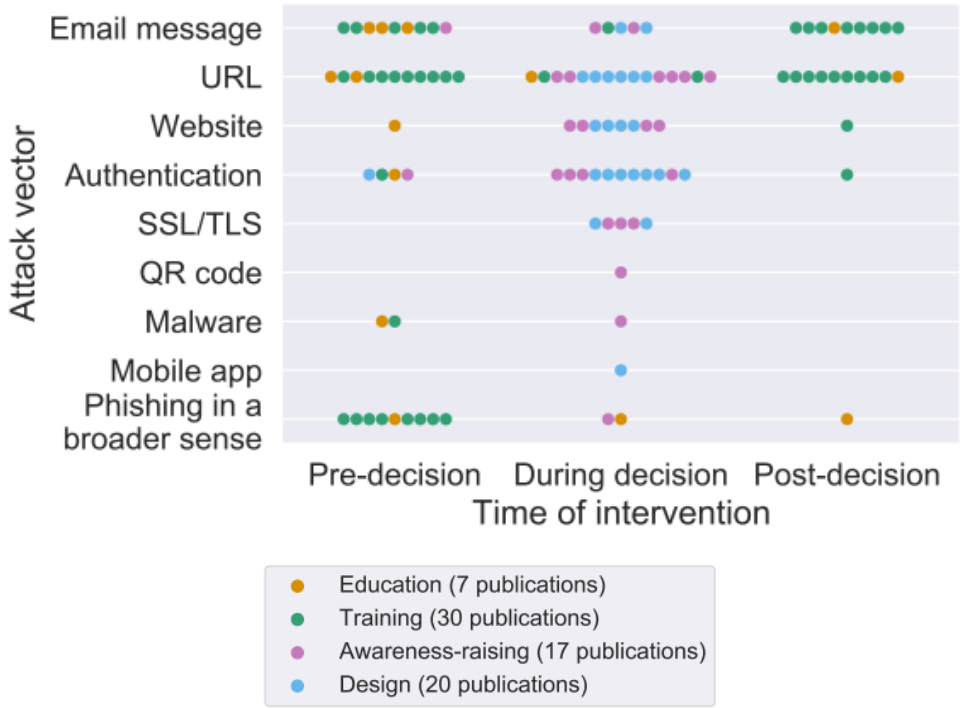


Figure 2-5 Overview of user-oriented phishing intervention literature, spanned by attack vector, time of intervention, and intervention category

2.4.1 User Effort and Intervention Intrusiveness

One particularly salient finding is that most user-oriented phishing interventions encumber the user with additional effort with respect to their workload and time, for example, in the form of playing a training game (Sheng et al., 2007), interacting with embedded training (Kumaraguru et al., 2009), answering security questions (Wu et al., 2006b), or waiting for delayed link activation (Volkamer et al., 2017). Those seconds or minutes required to interact with an intervention cumulatively drain time from individual and organizational productivity.

Moreover, they often intrusively disrupt the user in their primary goals, hence again substantially decreasing productivity by distraction and potentially leading to stress and frustration. This aligns with Sasse (2015)'s observation that user time and effort are rarely at the forefront of security studies and that the issue of user effort and intrusiveness has scarcely been considered. Sasse has argued that designers of security tasks should focus on *"causing minimum friction"* and *"must acknowledge and support human capabilities and limitations"* (Sasse, 2015, p. 82). She has called for subjecting security measures to a cost-benefit test and to give up on perfection and focus on essentials. On the other hand, passive, that is, less intrusive interventions have been observed to be of limited success as of yet (Iacono et al., 2014; Kirlappos and Sasse, 2011; Schechter et al., 2007; Wu et al., 2006a). It hence remains the most challenging task to design effective user-oriented phishing interventions that prove themselves usable in individuals' everyday online activities, particularly with regard to user effort and intrusiveness. Digital nudging (Thaler and Sunstein, 2008; Weinmann et al., 2016) might constitute an unintrusive yet promising approach for this endeavor. In section 0, we evaluate which elements of prior, effective interventions could be classified as nudges retrospectively and present ideas for future approaches. As for training and education interventions, Cranor & Garfinkel have argued that "*the world's future cyber-security depends upon the deployment of security technology that can be broadly used by untrained computer users"*, hence questioning the usability of such approaches. It is still an open question whether interventions need to be understood by the user (e.g., via providing educational information) in order to be effective (Zimmermann and Renaud, 2021; Egelman et al., 2008), whereas it has been observed that intervention clearness (e.g., with regard to their message concreteness (Schuetz et al., 2020) or their location (Petelka et al., 2019)) increases effectiveness. This spans an interesting research area with potentially crucial insights for the design of future phishing interventions.

2.4.2 Digital Nudges as Phishing Interventions?

The concept of digital nudging has scarcely been drawn on in phishing intervention research as of yet. The term nudging has been introduced by Thaler & Sunstein (Thaler and Sunstein, 2008) in 2008. Digital nudges describe user-interface design elements that target automatic cognitive processes, such as biases or heuristics, to gently push end-users, with little mental effort, to perform the "right" behavior without limiting their choice set (Thaler and Sunstein, 2008; Weinmann et al., 2016). In this section, we aim to discuss the potential of digital nudges in phishing intervention research, especially since prior research in related fields, such as digital

privacy-protection or security choices (Zimmermann and Renaud, 2021; Renaud and Zimmermann, 2019; Choe et al., 2013), can serve as a solid basis to start from. Surprisingly, phishing intervention literature from 2018 onward has focused on education, training, and awareness-raising measures, while neglecting design interventions (see Table 2-3 in the appendix). Design phishing interventions might provide significant value to users' security if they succeed in nudging users towards secure behavior, while not being perceived as intrusive with regard to their primary goals.

In an extensive review, Caraban et al. (2019) have classified six distinct nudge categories in the area of human-computer interactions. In the following, we exemplarily discuss how existing interventions make use of several of those mechanisms already (although not labeling them as nudging) and present novel ideas on how digital nudging could be applied in future phishing intervention research.

Facilitate. Facilitating nudges use mechanisms to lessen users' effort. In our sample, highlighting domains (Lin et al., 2011; Volkamer et al., 2017) or sender addresses (Egelman et al., 2008) falls in this category since it makes it easier for users to spot the relevant part of an URL or email sender. We propose to take this approach further, for example, by displaying a link's domain next to the link text in an email, with only the domain being clickable.

Confront. Confronting nudges aim to create friction by throttling users' mindless activity or reminding them of the consequences. Several of the interventions in our sample can be described as such, for example, interactive awareness-raising measures as described in section 2.3.2. As we have argued in the previous section, burdening the user with intrusive distraction and effort cannot be an efficient answer to current and future challenges in cyber security. We hence argue that confronting nudges should be designed to be of minimal possible friction. For example, they could remind the user of consequences by making security risks tangible.

Deceive. Deceptive nudges influence the perception of the available options, e.g., by adding inferior alternatives or placebos. None of the analyzed interventions could be sorted into this category, and we do not deem deceptive nudges suitable for phishing intervention research.

Social Influence. This type of nudge makes use of social influences on people's choices. Examples of social influence within the analyzed articles include the comparison of facts and stories provided by peers vs. experts on anti-phishing education (Marsden et al., 2020) as well as Nicholson et al. (2017)'s social saliency nudge. Furthermore, social influence has been studied in a social learning environment in terms of gamified elements such as levels or leader boards (Silic and Lowry, 2020). Future social influence nudges could provide users with

information on, e.g., their vs. their peers' performance in phishing simulations or incident reporting activities.

Fear. Two research works of our sample (Schuetz et al., 2020; Jansen and Schaik, 2019) have introduced fear appeals as phishing interventions with promising results regarding users' protection motivation, attitudes, intentions and compliant behavior. However, both articles have studied fear appeals far from a real-world scenario, using text-based treatments and a survey instrument. We suggest that fear nudges, which, integrated in the user's course of action, aim to invoke fear to encourage a certain choice, are of high interest for future research. Nevertheless, they require ethical considerations (Renaud and Dupuis, 2019). As an example, we imagine a brief but concrete (Schuetz et al., 2020) and strong (Jansen and Schaik, 2019) fear appeal next to email attachments, addressing the risk in terms of financial losses and operational damage coming along with this file type and a potential malware infection. The fear appeal could be framed positively to address ethical concerns by showing how the user could protect against these threats easily.

Reinforce. Reinforcing nudges aim to support certain behaviors, e.g., by ambient feedback or just-in-time prompts. Regarding the first, we found mechanisms ranging from color-coding security indicators on websites (Herzberg and Jbara, 2008; Wu et al., 2006a) to providing customized background images (Schechter et al., 2007; Marforio et al., 2016) in our sample. One shortcoming of these interventions seems to be that users cannot distinguish between legitimate security indicators (such as a color code) and untrustworthy signs, such as arbitrary logos and certifications (Kirlappos and Sasse, 2011). One way to battle this could be to make ambient feedback more comprehensive or standardized, e.g., by color-coding complete email or website windows. Concerning just-in-time prompts, in order to condense prior warning interventions to the pure form of a digital reinforcement nudge, we ideate an authentication intervention that displays the domain of a login website above any login form when placing the cursor in the login field.

Finally, suitable nudges could be easily combined with other interventions types, for example, educational elements (Zimmermann and Renaud, 2021), as shown by successful examples (Yang et al., 2017; Schuetz et al., 2020; Jansen and Schaik, 2019). As illustrated in Table 2-3 in the appendix, interventions that combine educational with awareness-raising or design approaches have rarely been studied in phishing research as of yet.

2.4.3 Shifting Users' Cognitive Frame

From a different perspective, Wash (2020) has adduced IT experts' approach towards identifying phishing emails and has observed that experts naturally follow a three-stage process: (1) making sense of the email, relating it to one's personal context, and deriving required action (2) becoming suspicious and investigating, and (3) dealing with the email by deleting or reporting it. He argues that shifting the user's cognitive frame from sensemaking to investigation is crucial for the success of phishing prevention measures. However, half of the interventions in our literature sample have addressed training or education measures (see Figure 2-5). Those mostly neglect the initial process of noticing slight discrepancies or cues in an email in the sensemaking frame and provide support only in the investigation frame (e.g., how to analyze an URL). While Jensen et al. (2017b)'s mindfulness-based training aims to support users in their awareness of context, and such during their sensemaking process, long-term efficacy is uncertain.

At the same time, users' own security goals should not be neglected: Kirlappos and Sasse (2011) have argued that users do not focus on security warnings, but rather look for signs to confirm a website's trustworthiness. For example, users have been shown to trust websites that display advertisements affiliated with known entities or those with familiar website layouts - while both factors do not give evidence of the website's trustworthiness. Therefore, the authors have called for security education to consider the drivers of users' behavior in their respective situation and, conversely, to eliminate users' misconceptions that lead to insecure behavior.

We hence argue that future phishing interventions should strive to meet the user in their own respective sensemaking process, for example, when reading emails, shopping, or doing bank transactions online. Digital nudges might play an important role in this particular case, as well. Supporting the user's cognitive frameshift from the stage of sensemaking to the stage of investigating if certain cues or discrepancies are present will be an important path for future research and will complement the diverse landscape of education and training measures.

2.4.4 What About Malware?

Regarding the phishing attack vectors addressed across our literature data, we have found that more than half of the interventions focus on the attack vector URL, for example, by training users' skills in analyzing a link or raising their awareness in situ. Interventions supporting the user with deceptive email messages, disguised websites, and fraudulent authentication forms follow by far (see Figure 2-3).

Malware poses a tremendous risk through current cyber attack patterns (Patsakis and Chrysanthou, 2020; Cofense, 2020). Those attacks are often delivered by archive files or Microsoft Office documents which mimic, e.g., legitimate invoices. Since the user needs to download and open these files on their system, this presents quite a different attack procedure compared with clicking a link. Therefore, it is striking that only three publications have included educational, training, or awareness-raising interventions in their works that address malware alongside other attack vectors. None of the articles in our sample has focused on studying interventions that primarily support users in detecting or handling malware, nor have the challenges of malware interventions compared to previous phishing intervention research been addressed. We therefore strongly suggest further research to expand previous approaches on phishing interventions in terms of the attack vector by taking into account malicious files and developing interventions that address the actual threat landscape.

2.4.5 Tailored Interventions

In the context of user interventions in cyber security, several studies have pointed out the potential of personalization regarding user traits (Egelman and Peer, 2015; Jeske et al., 2014; Peer et al., 2020), or the importance of context (e.g., personal vs. organizational (Schuetz et al., 2020)). It has been argued that using tailored instead of one-size-fits-all interventions may enhance their efficacy and user compliance (Egelman and Peer, 2015).

Interestingly, our literature review does not reveal a strong focus on tailored user interventions to prevent phishing attacks. However, some of the approaches were indeed implemented for specific target groups mainly for rather heterogeneous groups of employees (Silic and Lowry, 2020), or children (Lastdrager et al., 2017). Since spear phishing attacks are specifically targeted at personal or contextual vulnerabilities, considering users' traits, capabilities and requirements when developing and evaluating user interventions may be a decisive factor for their efficacy, suggesting a scope for future research.

2.4.6 Methodological Aspects

As described in section 2.3.1, current research often lacks realism regarding the experimental setup since it remains challenging to study a phenomenon of deception that usually takes place during users' secondary tasks. Therefore, we argue that future research should not only focus on designing user-oriented phishing interventions, but also on developing experimental setups that account for a realistic analysis of users' security behavior.

Furthermore, we have found that the effect of recurring interventions has been studied scarcely (see section 2.3.4). However, many interventions in our sample are designed to train, warn or guide users recurrently. Factors such as habituation (Vance et al., 2018) or security fatigue hence could have important effects. This proves another major shortcoming in prior phishing intervention research, which should be considered by future works.

2.4.7 Limitations

In this work, we have carefully selected (usable) security-specific databases to include a large number and variety of publications. Furthermore, the chosen search term was rather broad, and additional sources (such as security conferences) were considered to avoid overlooking relevant findings. Nevertheless, the list of publications analyzed in this research is probably not exhaustive. Furthermore, the features of the different phishing interventions were described in varying detail due to the individual focus and comprehensiveness of the articles. It is thus possible that certain interventions were classified differently by us than the authors themselves would have classified them. Therefore, this systematization of knowledge does not serve as an endpoint but as a starting point for identifying the current state, potential research gaps, and relevant paths for future work. We hope to not only provide a relevant summary and systematization of existing strategies for usable security-related researchers and practitioners, but especially to encourage future studies in this increasingly relevant domain, where the human factor plays an essential role.

2.5 Conclusion

Phishing does not cease to be a threat to both personal and organizational data and operational security. It directly targets the human factor via deceptive emails, attachments, and websites, hence calling for user-oriented interventions that support individuals in recognizing and fending off such attacks. In this work, we have systematically analyzed 64 phishing intervention research articles for methodology, intervention type, attack vector, intervention time and user interaction, and have derived a taxonomy of user-oriented phishing interventions. Connecting the findings across the dimensions of analysis, as well as considering current movements in usable security research, we have revealed relevant insights and potential avenues for future work. The latter can be summarized as follows:

Minimize user effort and intervention intrusiveness. How can we design effective phishing interventions that cause minimum friction with the user's course of action and do not cumulatively burden the user with secondary time and workload? Which role does educational

information play in intervention effectiveness, compared with intervention clearness and concreteness?

Explore the potential of digital nudging. How can facilitating, confronting, reinforcing, fear, or social influence nudges support users' course of action with regard to secure online behavior?

Help users shift their cognitive frame. How can we support users in the cognitive process of shifting from their primary goal of sensemaking towards noticing discrepancies if "something is off"? How can we transfer experts' expertise with phishing detection into effective end-user interventions?

Protect users from malware attacks. Which kinds of interventions can help to protect users from malware attacks? Which novel challenges do arise for malware-focused interventions, compared with threats employing malicious URLs or websites?

Explore tailored interventions. How can tailored phishing interventions enhance previous approaches?

Develop realistic experimental setups and study long-term effects. Which novel ways can be employed to align experimental setups with the nature of phishing and to account for longitudinal effects?

With this article, we hope to provide a comprehensive starting point as well as inspiration for future user-oriented phishing intervention research.

2.6 Acknowledgements

This research work has been funded by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

2.7 Appendix

Database	After search	After exclusion
ACM	270	35
IEEE	869	15
Web of Science	970	25
NDSS/(Euro)USEC	5	2
USENIX Security/SOUPS	8	3
Other	2	2

Table 2-2. Number of articles included in the literature review before and after applying the exclusion criteria

	Sample Size	Lab Study	Online Study	Field Study	Survey	Conceptual	Education	Training	Awareness-raising	Design	Email	URL	Website	Authentication	SSL	Other	Pre-Decision	During Decision	Post-Decision	Interactive	Passive	Educational	Non-Educational
Author	**Method**						**Intervention Category**				**Attack Vector**						**Time of Intervention**			**Activity**		**Education**	
(Abbasi et al., 2016)	509		•						•			•	•					•		•			•
(Alnajim and Munro, 2009)	36	•						•				•	•						•	•		•	
(Arachchilage et al., 2016)	20	•						•									•			•		•	
(Baslyman and Chiasson, 2016)	21	•						•									•			•		•	
(Beguin et al., 2019)	14	•						•									•			•		•	
(Blythe et al., 2011)	/					•	•		•								•	•		•		•	
(Burns et al., 2019b)	400			•				•			•	•							•	•		•	
(Burns et al., 2013)	/					•		•									•			•		•	
(Canova et al., 2015)	19	•	•					•				•					•			•		•	
(Caputo et al., 2013)	1,359			•				•			•	•							•	•		•	
(Carella et al., 2017)	150			•				•									•			•		•	
(Cuchta et al., 2019)	4,777			•				•			•	•							•	•		•	
(De Ryck et al., 2013)	/					•				•			•					•			•		•
(Dhamija and Tygar, 2005)	/					•				•				•				•			•		•
(Egelman et al., 2008)	60	•							•				•					•		•	•		•
(Fatima et al., 2019)	63	•						•									•			•		•	
(Gastellier-Prevost et al., 2011)	/					•			•				•					•		•			•
(Gokul et al., 2018)	8,071		•					•				•					•			•		•	
(Greene et al., 2018b)	ca. 70			•	•			•			•								•		•		•
(Hale et al., 2015)	/					•		•									•			•		•	
(Hale and Gamble, 2014)	/					•		•									•			•		•	
(Herzberg and Jbara, 2008)	23	•								•		•			•			•			•		•
(Herzberg and Margulies, 2013)	400			•						•				•				•		•			•
(Iacono et al., 2014)	18		•							•			•					•			•		•
(Jakobsson and Myers, 2007)	/					•				•				•				•			•		•
(Jansen and Schaik, 2019)	786				•		•		•					•			•				•	•	
(Jensen et al., 2017b)	355			•				•			•	•					•			•		•	
(Kirlappos and Sasse, 2011)	36		•							•			•					•		•		•	

	Sample Size	Lab Study	Online Study	Field Study	Survey	Conceptual	Education	Training	Awareness-raising	Design	Email	URL	Website	Authentication	SSL	Other	Pre-Decision	During Decision	Post-Decision	Interactive	Passive	Educational	Non-Educational
(Kumaraguru et al., 2010)	4,517			•				•			•	•							•	•		•	
(Kumaraguru et al., 2009)	515		•					•			•	•							•	•		•	
(Kumaraguru et al., 2008)	311			•				•			•	•							•	•		•	
(Kumaraguru et al., 2007)	30	•						•			•	•							•		•	•	
(Lastdrager et al., 2017)	353	•					•				•	•	•				•			•		•	
(Li et al., 2012)	20	•								•		•	•					•		?	?	?	?
(Lin et al., 2011)	22	•								•		•						•			•		•
(Marforio et al., 2016)	221	•								•								•			•		•
(Marsden et al., 2020)	11,968		•					•				•							•	•		•	
(Miyamoto et al., 2014)	23	•								•		•		•				•		•			•
(Nicholson et al., 2017)	279		•							•	•							•			•		•
(Perrault, 2018)	462				•			•			•	•					•			•		•	
(Petelka et al., 2019)	701		•						•			•						•		•			•
(Reeder et al., 2018)	773			•					•					•	•	•		•		•		?	?
(Reinheimer et al., 2020)	409			•			•					•				•	•			•		•	
(Ronda et al., 2008)	2,050			•						•				•				•			•	•	
(Schechter et al., 2007)	67	•							•	•		•	•	•				•		•			•
(Schuetz et al., 2020)	264		•				•		•								•			•		•	
(Scott et al., 2014)	/					•		•				•					•			•		•	
(Sheng et al., 2007)	42	•						•				•					•			•		•	
(Silic and Lowry, 2020)	384			•				•			•	•					•			•		•	
(Stembert et al., 2015)	24	•						•	•		•	•						•		•	•	•	
(Stockhardt et al., 2016)	81	•						•				•					•			•		•	
(Varshney et al., 2012)	/					•				•				•				•			•		•
(Volkamer et al., 2017)	16			•					•	•		•						•		•		•	
(Wash and Cooper, 2018)	1,945			•			•	•			•	•							•	•		•	
(Weanquoi et al., 2017)	/					•		•									•			•		•	
(Wen et al., 2019)	39	•						•			•	•				•	•			•		•	
(Wiese et al., 2018)	18		•						•	•	•							•		•		•	
(Wu et al., 2006a)	21	•							•	•				•				•		•		•	

	Sample Size	Lab Study	Online Study	Field Study	Survey	Conceptual	Education	Training	Awareness-raising	Design	Email	URL	Website	Authentication	SSL	Other	Pre-Decision	During Decision	Post-Decision	Interactive	Passive	Educational	Non-Educational
(Wu et al., 2006b)	30	•							•	•		•			•			•			•		•
(Xiong et al., 2019)	639		•					•				•						•		•		•	
(Yang et al., 2017)	63			•			•		•			•						•		•		•	
(Yao and Shin, 2013)	20	•							•							•		•		•		•	
(Yee and Sitaker, 2006)	/					•				•				•				•		•		•	
(Yue, 2012)	/					•			•					•				•		•			•
Sum (N=64)		20	12	16	3	13	7	31	17	20	17	33	10	12	4	4	23	31	11	48	16	43	19

Table 2-3. Results of the literature review, sorted alphabetically by first author

Chapter 3: Underlying Mechanisms of Interdependent Privacy Decision-making

Title:	Exploring Interdependent Privacy – Empirical Insights into Users' Protection of Others' Privacy on Online Platforms
Authors:	Anjuli Franz, Technische Universität Darmstadt, Germany Alexander Benlian, Technische Universität Darmstadt, Germany
Published in:	Electronic Markets (2022) (forthcoming, spring 2023)

Abstract: Recent information privacy research has started to spark a debate about privacy infringements that happen not on an individual, but on a multi-party level. Here, a person's own information privacy is affected by the decisions of others – a phenomenon referred to as interdependent privacy. Building on the 3R Interdependent Privacy Protection Framework, we explore the underlying mechanisms of how and why interdependent privacy violations happen and how they can be remedied. Drawing on an online vignette experiment (N = 330), we investigate the efficacy of an interdependent privacy salience nudge and reveal that it can decrease the likelihood that users disclose others' personal information by 62%. Furthermore, we develop a novel measurement instrument and empirically validate that users' decision to disclose others' personal information to an online platform is formed via a serial mediation mechanism through users' realization of the data transfer, recognition of others' ownership, and respect for others' rights. We discuss important implications for both theory and practice.

Keywords: Interdependent privacy, Peer disclosure, Online platforms, Privacy nudge, Online vignette study, Serial multiple mediation

3.1 Introduction

Privacy issues challenge researchers and regulators because of their immense complexity, and have been discussed through various lenses. Modern perspectives have matured from viewing privacy as a transactional process of information disclosure, to viewing it as a multi-faceted, socially constructed phenomenon that is closely tied to real-world modern networked technologies, and that we should endeavor to embed in the design of the tools and services we

use daily (Bélanger and James, 2020; Knijnenburg et al., 2022). Particularly in the context of online platforms, where personal data is being generated and shared at lightning speed, privacy losses and violations are far from trivial to perceive and decide upon, and often remain unconsidered (Lowry et al., 2017; Garcia, 2017).

When investigating privacy concerns or disclosure decisions, the preponderance of privacy literature has limited its scope to a dyadic understanding of privacy, e.g., a dyadic information transfer between a company and an individual (Kamleitner and Mitchell, 2019). In contrast, recent research has called for a more versatile multi-level understanding of privacy to be able to explore complex disclosure decisions in progressively sophisticated digital environments (Bélanger and James, 2020). One crucial factor that makes privacy a highly complex affair are the various types of inherent connections among individuals. Since human beings are socially embedded and bond with each other by exchanging personal information, their personal data is often not only owned by themselves, but also co-owned by others. For example, chances are high that there are hundreds of co-owners of your phone number and email address (e.g., friends who have stored your contact information in their address book), and that a social platform (e.g., LinkedIn) has collected various information on your interests and preferences. This makes privacy protection an interdependent phenomenon (e.g., Biczók and Chia, 2013; Cao et al., 2018; Wirth et al., 2019) since the violation of an individual's privacy rights can happen through others, potentially without the original owner even noticing.

In recent years, several research streams have approached this phenomenon employing, for example, economic models (e.g., Cao et al., 2018; Symeonidis et al., 2018) or empirical studies on users' behavior (e.g., Olteanu et al., 2018; Pu and Grossklags, 2017; Pu and Grossklags, 2015), have analyzed legal aspects (e.g., Symeonidis et al., 2018) or developed conceptual frameworks (e.g., Jia and Xu, 2016; Kamleitner and Mitchell, 2019). Among the latter, Kamleitner and Mitchell (2019) have proposed the "3R Interdependent Privacy Protection Framework", which postulates a sequential chain of the underlying mechanisms realization of the data transfer (RE), recognition of others' ownership (RC), and respect for others' rights (RS) forming an individual's decision to protect others' personal data. While the 3R framework advances our understanding of how interdependent privacy decision-making might unfold and can inform future research, it has not been empirically validated to date. Furthermore, while information privacy research on the individual level has investigated interventions such as transparency tools to help users make informed decisions (e.g., Almuhimedi et al., 2015; Wang et al., 2014), we have little knowledge on effective countermeasures that can help to mitigate

interdependent privacy violations. This is reflected in current regulatory efforts to protect individuals' and third parties' privacy, for example, the GDPR (European Parliament and Council, 2016): Whereas the GDPR has achieved major improvements to protect users' own information (e.g., mandatory opt-in mechanisms when collecting users' data for marketing purposes), little has been done to protect users from interdependent privacy violations by their peers. Since others' decisions on our privacy can have significant impacts on our everyday lives, we argue that it is paramount that we (1) understand the underlying mechanisms of interdependent privacy violations, and (2) find effective remedies that can serve as design suggestions for novel regulatory strategies. In this work, we therefore raise the following research questions:

1: To what extent can the 3R mechanisms underlying users' interdependent privacy decision-making be empirically validated?

2: How can interdependent privacy infringements be reduced via design choices, such as an interdependent privacy salience nudge?

To answer our research questions, we draw on the theoretical lens of the "3R Interdependent Privacy Protection Framework" established by Kamleitner and Mitchell (2019). We conduct a quantitative vignette-based online experiment with N = 330 Instagram users, motivated by an actual Instagram prompt that encourages users to violate others' privacy. In our experiment, we investigate the effect of an interdependent privacy salience nudge that aims to increase the salience of the other in the data transfer. We analyze our experimental data employing a serial multiple mediation analysis, which supports our hypotheses. Our post-hoc analysis of qualitative statements gives richer insights into participants' motives, and further confirms our theoretical model.

Our study contributes to research on interdependent privacy in several important ways. First, we investigate the effect of an interdependent privacy salience nudge and show that it can significantly improve users' protection of their peers' privacy online. Second, we empirically evaluate the "3R Interdependent Privacy Protection Framework" (Kamleitner and Mitchell, 2019). Our results indicate a three-stage mediation of the effect of our interdependent privacy salience nudge on users' disclosure of others' information through RE, RC, and RS, which validates Kamleitner and Mitchell (2019)'s theoretical model of users' interdependent privacy decision-making. Lastly, as part of our study, we develop and validate a measurement instrument for RE, RC, and RS, which can be useful for future research in this field. Our work implicates valuable insights for regulators, as it can serve as a starting point for overcoming

current policy inadequacies (e.g., in the GDPR) with regard to interdependent privacy infringements.

The remainder of this paper is organized as follows. In section 3.2, we first introduce the phenomenon of interdependent privacy in the context of social platforms by giving several real-world examples as well as a brief overview of pertinent literature. We then introduce Kamleitner and Mitchell (2019)'s "3R Interdependent Privacy Protection Framework" as a theoretical lens for our study. Lastly, we turn to the concept of digital nudging in privacy, hence laying the conceptual foundation for the interdependent privacy salience nudge employed in our experiment. In section 3.3, we then develop our research model and hypotheses. We proceed with describing our research methodology and introducing the concept of serial mediation in section 3.4. After presenting the quantitative and qualitative results of our empirical study in section 3.5, we discuss our findings as well as our contributions to theory and practice in section 3.6. Finally, in section 3.7, we summarize the findings of this article.

3.2 Theoretical Background

3.2.1 Privacy Interdependence

In 1970, long before mobile devices and social networking platforms have emerged as omnipresent parts of our lives, Westin (1970) has defined privacy as "the ability to control, edit, manage, and delete information" about oneself and to "decide when, how, and to what extent information is communicated to others" (p. 7). Since then, privacy has arisen to be one of the most crucial concepts of our time: While personal information (e.g., photos, preferences or location data) is being generated and shared online at a rapid pace, recent sociopolitical movements (e.g., Harwell and Harris, 2021; Isaak and Hanna, 2018) demonstrate why privacy rights are of paramount importance for individuals' freedom and sovereignty. In 2018, based on the concept of privacy as a fundamental human right, the European Union (EU) has issued the General Data Protection Regulation (GDPR) (European Parliament and Council, 2016). The GDPR regulates the processing of personal data related to citizens of the EU and has acted as a catalyst for major transformations of privacy policies worldwide (Li et al., 2019; Linden et al., 2020). In an online context, however, privacy losses or violations represent intricate problems for both users and regulators, since they are often nontrivial to perceive and decide upon (Garcia, 2017). Contrarily, privacy is a highly complex affair, with one crucial factor being the various types of inherent connections among human beings and their personal data (Biczók et

al., 2021). In the following, we will illustrate this interconnectedness drawing on the example of online platforms.

Imagine a purely individualistic perspective, where a person's own privacy is affected only by their own decisions. Here, common theoretical approaches such as the privacy calculus model (e.g., Dienlin and Metzger, 2016; Dinev and Hart, 2006; Kehr et al., 2015) can give insight into users' analysis of perceived costs (e.g., privacy risks) and benefits (e.g., entertainment), and hence the formation of their intention to disclose their own information. With respect to online platforms, such as Facebook or LinkedIn, this assumption would hold only if individuals used such platforms in isolation. This is, however, not the case: For many online platforms, the interconnectedness of their users' data lies at the core of their business. For example, when a user installs a third-party application on Facebook, the application might collect not only a focal user's, but also their friends' personal information (Symeonidis et al., 2018). This has laid the foundation for the Cambridge Analytica scandal, which came to light in 2018 (Isaak and Hanna, 2018): While only 270.000 users installed the company's app-based personality quiz on Facebook, Cambridge Analytica harvested the personal data of an estimated 87 million people and used it for micro-targeting during the 2016 US election campaign (Kamleitner and Sotoudeh, 2019). As a second example, LinkedIn, a professional social network, relies on users' opinions on their contacts' skills (e.g., "Help us identify Anna Smith's top skill") in order to offer and sell personalized job opportunities. These examples demonstrate that a person's own privacy is not only affected by their own decisions, but is also controlled by the actions of other individuals or organizations. We refer to this phenomenon as interdependent privacy, where "personal information is shared without the knowledge and/or direct consent of the data subject" (Biczók et al., 2021). The notion of privacy interdependence renders the aforementioned perspective of an individual privacy calculus obsolete.

In recent years, researchers from various fields (such as information security, information systems, economics or marketing) have started to spark a debate of the consequences, risks and potential mitigations of privacy interdependence. Reviewing recent literature, we found that one central concept is users' awareness of interdependent privacy risks (e.g., Biczók and Chia, 2013; Symeonidis et al., 2018): While, in an analog world, interdependent privacy protection seems to work according to implicitly negotiated "norms about what, why, and to whom information is shared within specific relationships" (Martin, 2016, p. 551), these negotiations appear to be largely absent when we consider interdependent privacy in an online context (Kamleitner and Mitchell, 2019). Prior research has demonstrated across data types (e.g.,

contact information or photos) that users are less considerate towards the privacy of their peers, compared to their own (Marsch et al., 2021). At the same time, new information and communication technologies allow for a tremendously larger scope of potential interdependent privacy violations, since users are able to automatically and effortlessly collect and disclose others' information. To illustrate this, we borrow from a fictive scenario introduced by Kamleitner and Sotoudeh (2019), and imagine a person called Ada, who is on a trip to explore a foreign city. Ada is looking for a nice place to stay, and asks a woman passing by if she has any tips. The woman responds: "Well I do have some really good recommendations, but first give me the name and phone number of your father, and maybe also a picture of him." Ada is baffled, refuses, and walks away. She implicitly feels that this information is personal, and not hers to share. In an online setting, however, Ada would consult a travel booking app, with hundreds of accommodation options being just one click away. The app might ask for access to her contacts. Her contact list includes information (such as a name, phone number, picture, and birthday) on her father, as well as pretty much anyone Ada knows. Yet, she might simply click "Allow Access", hence becoming a sharer of her contacts' data to an online platform without her contacts even knowing about it.

Presently, the issue of interdependent privacy displays a regulatory loophole for the GDPR (Kamleitner and Sotoudeh, 2019). The GDPR limits its scope to a dyadic understanding of privacy (e.g., between a company and a consumer), while leaving room for gray area with regard to interdependent privacy infringements. It specifies informed consent by the original data subject as a lawful prerequisite for the processing of personal data (Art. 6, GDPR), and further specifies that the original owner needs to be notified and provided with easy withdrawal of consent (Art. 7, GDPR). This regulation assumes that it is always clear who the original owner of personal information is. However, whereas Ada gives consent to share her contacts' data, her contacts might claim the ownership and privacy rights towards this information. While the GDPR specifically excludes the processing of personal information for household or purely personal purposes (Art. 2, GDPR), it is questionable if this exception covers the transfer of personal information of several hundred individuals to a company, such as an online platform, that processes this information as part of its business model. The negligence of interdependent privacy phenomena hence poses a major shortcoming of the GDPR in its current version.

Previous literature has approached the concept of interdependent privacy from various angles. In a recent meta-analysis, Humbert et al. (2019) have summarized and analyzed prior works across the research landscape. While "interdependent privacy" seems to be the most widely

used term, a variety of different terminologies is being used, such as collective privacy (e.g., Squicciarini et al., 2009), multiparty privacy (e.g., Thomas et al., 2010), or peer disclosure (e.g., Cao et al., 2018; Chen et al., 2015a). Several researchers have employed game-theoretical models to investigate the externalities of privacy interdependence (e.g., Biczók and Chia, 2013; Cao et al., 2018). For example, Symeonidis et al. (2018) have calculated the extent of collateral information collection by third-party apps on Facebook, finding that a user's chance of having their personal data shared with third-party apps through their friends is greater than 80%. This enables practices such as shadow profiling, where a company composes profiles of individuals based on data gathered from other users on a large scale (Garcia, 2017). Other works have focused on empirically exploring interdependent privacy behavior, for example, by investigating the monetary value which users of online services place on their contacts' personal information (Marsch et al., 2021; Pu and Grossklags, 2015), or by analyzing the roles of information sensitivity (Wirth et al., 2019) or sharers' anonymity (Pu and Grossklags, 2017).

Whereas previous research has yielded important insights into the topic of privacy interdependence, we have only little knowledge on the how and why, that is, on the underlying mechanisms of interdependent privacy behavior. By mechanisms, we refer to social mechanisms that act as "building blocks for the construction of causal explanations of social phenomena" (Avgerou, 2013, p. 407), which drive the process of forming an interdependent privacy decision and explain the observed behavior. One approach to tackle the how and why of interdependent privacy behavior is Kamleitner and Mitchell (2019)'s conceptual "3R Interdependent Privacy Protection Framework", which we will introduce in the following section.

3.2.2 The 3R Interdependent Privacy Protection Framework

In their framework, Kamleitner and Mitchell (2019) have approached the phenomenon of interdependent privacy infringements by drawing on the conceptual commonality between personal data and property. Individuals feel a sense of ownership for property, and the protection of such property necessitates the "cooperation of others and their respect of what is 'ours'" (Kamleitner and Sotoudeh, 2019, p. 2). While individuals also feel a sense of ownership for personal information, property and personal information differ with regards to their tangibility. Property refers to the right to one's possession, that is, to goods that are mostly tangible. For example, a house can be touched and seen, and can be held by only one or few individuals at a time. We are hence usually aware that someone owns it. On the contrary, personal information is mostly intangible. Imagine, for example, a phone number. Since it can

be held by an unlimited amount of people at a time, it is practically impossible to oversee how often it has been shared. Moreover, while the transfer of property usually takes place via an active acquisition (for instance, buying a house), the transfer of personal information often arises as a side effect of our daily activities. For example, when using an online platform, personal information is being shared to other individuals or organizations without the transfer of data as a good being in the focus of attention. Kamleitner and Mitchell (2019) argue that these fundamental differences make it much easier to trespass on privacy, that is, the right to one's personal information, than property. While property infringements mostly arise from a failure of respect, interdependent privacy violations can be caused by failures at antecedent stages (Kamleitner and Mitchell, 2019; Kamleitner and Sotoudeh, 2019) have derived three sequential steps that users need to take in order to protect others' personal information online: realization of the data transfer (RE), recognition of others' ownership (RC), and lastly respect for others' rights (RS). Figure 3-1 illustrates these steps based on the introductory example of Ada downloading an app.

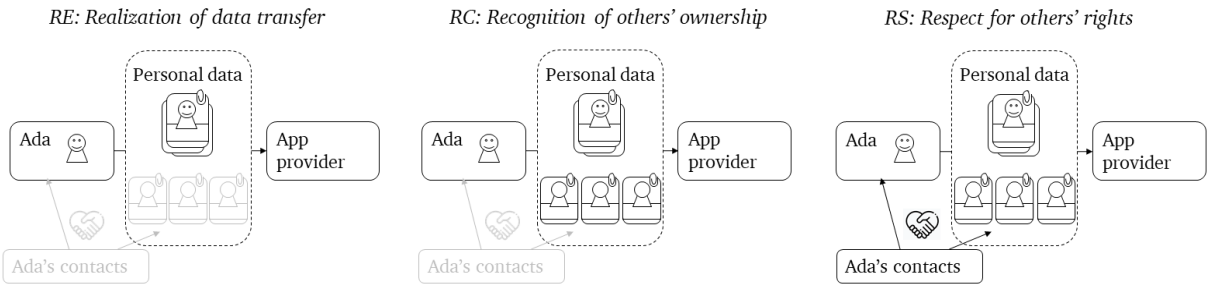


Figure 3-1. The 3R Interdependent Privacy Protection Framework

According to the 3R framework, users' realization of the data transfer (RE) represents the first step toward protecting others' personal information. Imagine a sharer synchronizing their address book with the Instagram app. In order to realize that this implies transferring co-owned data, the sharer first needs to realize that they are about to transfer a good to another party at all. Users' RE is based on the presence of two conditions: The sharer first needs to overcome the intangible nature of information which makes it difficult to truly comprehend data as a good, and then must realize that this good is about to be transferred from one party to another. In our example, Ada might press "Allow Access" without realizing that this will transfer data from her phone to the app provider, which would, in this moment, leave her unable to recognize others' being involved in the data transfer.

Provided that a sharer realizes that they are about to transfer a good, they then need to recognize others' ownership[7] (RC) of this good. When the app asks Ada for access to "her" contacts, she might not even consider the possibility of others holding a stake. Furthermore, the feeling of self-entitlement might weaken her recognition of others' ownership: Ada might recognize that others are somewhat involved in the data about to be transferred, but might feel self-entitled to this data. This feeling of entitlement might arise, for example, if the sharer has self-collected the information on a device that they own (e.g., their phone), or if they are in close relationship to the other (e.g., a partner or parent). Both the visibility of the other and the recognition of others' entitlement are hence important prerequisites for users' RC.

Lastly, respect for others' rights (RS) presents the final stage to prevent interdependent privacy violations: Once the sharer has recognized that what they are about to share belongs to another person, their respect towards others' privacy rights affects their further actions. There are several options for a sharer to respect others' privacy, for example, by refraining from the data transfer at all, or by obtaining consent from the other. According to Kamleitner and Mitchell (2019), there are two main antecedent forces that play a role in users' formation of respect for others' rights. First, while, in an analog world, norms of respect for what belongs to others are implicitly negotiated, society seems to trivialize disrespect towards others' privacy in digital settings. Users might thus consider it socially acceptable to infringe on others' privacy, because "everyone does it". Second, users might weigh their own benefit of the interdependent privacy violation against their own or others' costs, and hence deliberately infringe on others' privacy by knowingly putting their own interests above those of others.

The 3R Interdependent Privacy Protection Framework hence postulates three sequential steps where RE is a prerequisite for RC, and RC in turn is a prerequisite for RS. Together, the three steps act as a mechanism for users' formation of an interdependent privacy decision.

3.2.3 Privacy Nudging

Since Thaler and Sunstein (2008) have introduced the concept of nudging in 2008, it has found widespread attention in both research and practice. Nudges describe design elements that target automatic cognitive processes, such as biases or heuristics, to gently push individuals to perform the "right" behavior without limiting their choice set. Examples from the analog world

[7] While Kamleitner and Mitchell (2019) refer to the second stage as "recognition of others' rights", we chose to use the term "recognition of others' ownership", since we think that it (1) better represents the underlying concept and (2) is more distinguishable from the third stage, "respect for others rights".

include default options in organ donation or speed signs displaying smiling or frowning emoji. In information privacy research, prior works have started to investigate the potential of digital nudges (Schneider et al., 2018) in persuading users to act in a privacy-preserving manner in individual privacy contexts (e.g., Acquisti et al., 2017; Almuhimedi et al., 2015; Wang et al., 2014). Furthermore, recent research has started to call for the design and evaluation of nudges and permission interfaces "that approach privacy not simply as an individual issue, but as an interdependent and collective concern" (Marsch et al., 2021, p. 17).

Reviewing the vast body of literature on nudging, we find that nudges can take on various designs. Popular mechanisms are, for example, default options, positioning or color coding, reminding of the consequences, or enabling social comparison (Caraban et al., 2019). In their paper on the 3R framework, Kamleitner and Mitchell (2019) have suggested several interventions to improve interdependent privacy protection across stakeholders, e.g., requiring additional steps of decision control in the transfer process, or a preview of the actual data which is about to be shared. These suggestions provide a valuable basis for the design of nudges in an interdependent privacy context.

3.3 Research Model and Hypothesis Development

Building upon prior works on privacy nudging (e.g., Almuhimedi et al., 2015; Wang et al., 2014; Zhang and Xu, 2016) and interdependent privacy protection (Kamleitner and Mitchell, 2019), we develop a research model which suggests that users' RE, RC, and RS carry over the effect of an interdependent privacy salience nudge (IPN) to users' decision to disclose others' information (DOI). Figure 3-2 depicts our proposed research model.

Prior research on privacy nudging suggests that nudges that are designed to enable informed decision-making can facilitate privacy-aware behavior (Almuhimedi et al., 2015; Wang et al., 2014). For instance, confronting users with feedback on how often their location data was shared with apps has been shown to make users control their app permissions more restrictively (Almuhimedi et al., 2015). Regarding the violation of interdependent privacy, we thus hypothesize the following:

H1: The presence (vs. absence) of an interdependent privacy salience nudge (IPN) decreases users' disclosure of others' information (DOI).

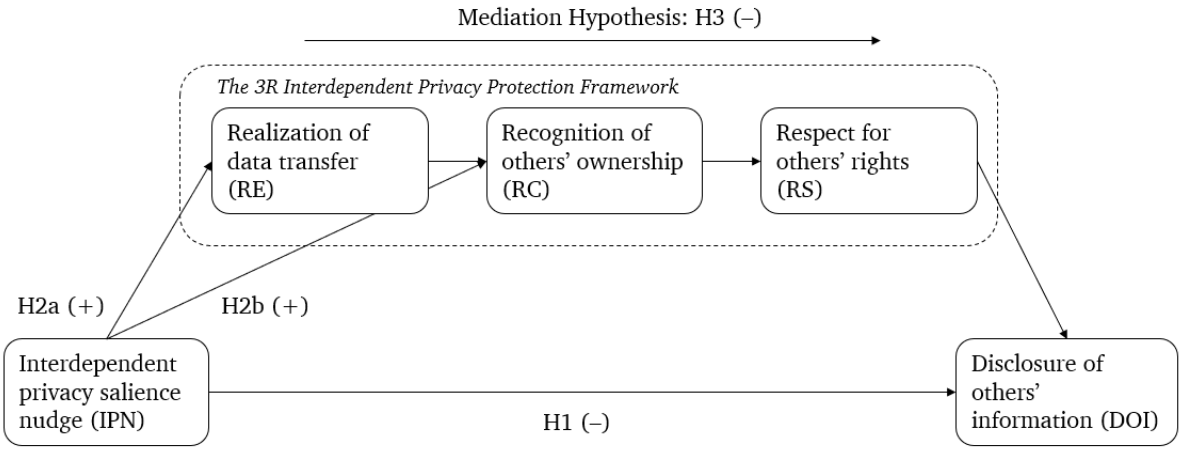


Figure 3-2. Research model

In our remaining hypotheses, we aim to dive deeper into the underlying mechanisms of this effect. We use (Avgerou, 2013)'s definition of social mechanisms as sequences of events unfolding in time, that generate causal processes and ultimately observed outcomes. Social mechanisms might show, for example, how individuals develop specific meanings of an information system, or why they act in a particular way when interacting with technology in a certain context. We draw on the 3R framework (Kamleitner and Mitchell, 2019), where the salience of the good (i.e., the personal data about to be transferred) and the salience of the transfer have been suggested as antecedents of users' realization of the data transfer. Prior research has demonstrated that providing users with salient and accessible privacy information guides users' attention towards the information disclosure and its potential risks (Tsai et al., 2011). Getting back to our example of Ada downloading an app, we argue that it would have been easier for her to realize the data transfer if the app would have provided her with transparent and detailed information instead of simply asking for access to her contacts. For an IPN that increases the salience of both the data and the transfer, we hypothesize the following:

H2a: The presence (vs. absence) of an interdependent privacy salience nudge (IPN) increases users' realization of the data transfer (RE).

Analogously, the salience of the other has been named as the antecedent of users' recognition of others' ownership when sharing information in an interdependent privacy context (Kamleitner and Mitchell, 2019). Increasing the salience of the other will hence increase the user's attention towards the role of others' ownership during the data transfer: If the app had explicitly informed Ada that she was about to share information that does not belong to her, but to others, she would have been more likely to recognize others' ownership of the data. For an IPN that increases the salience of the other, we hence posit:

H2b: The presence (vs. absence) of an interdependent privacy salience nudge (IPN) increases users' recognition of others' ownership (RC).

H2a and H2b reflect our expectation of how the salience and accessibility of the displayed interdependent privacy information affect users' RE and RC.

Drawing on the 3R Interdependent Privacy Protection Framework (Kamleitner and Mitchell, 2019), we predict that RE will feed into users' RC, which will in turn increase their RS, and ultimately decrease their DOI. In other words, we posit that the effect of the IPN on users' DOI takes place via a three-stage serial mediation through RE, RC, and RS:

H3: Users' realization of the data transfer (RE), recognition of others' ownership (RC), and respect for others' rights (RS) will act as a three-stage serial mediator for the effect of the interdependent privacy salience nudge (IPN) on users' disclosure of others' information (DOI).

3.4 Research Methodology

3.4.1 Experimental Design and Procedure

To test our hypotheses, we conducted an online experiment and embedded our treatments based on vignettes depicted in an online survey. The vignette methodology has been validated as an effective measurement technique for assessing users' perceptions of and responses to specific information privacy-related conditions (Benlian et al., 2020; Lowry et al., 2019; Warkentin et al., 2017). We chose the social networking platform Instagram as the context for our study for two main reasons. First, Instagram and its parent company, Meta Platforms (formerly Facebook), have been increasingly facilitating users' voluntary information disclosure about not only their own, but also others' information (Alsarkal et al., 2018; Symeonidis et al., 2018). Employing a vignette scenario on Instagram hence allowed us to use a real-world prompt which encourages interdependent privacy violations on online platforms. Second, Instagram is among the most popular social networks as of 2021 (Statista, 2021), which allowed for a large number of potential participants in our study.

In our online vignette experiment, participants were welcomed and told that they participated in a study on Instagram use. They were asked to answer all questions honestly, and were told that there were no right or wrong answers. Furthermore, they were informed about their anonymity and the intended use of the collected data. Participants where then asked to imagine that they were logged into their personal Instagram account. They were told to imagine that,

while browsing their Instagram feed, a prompt pops up, which was shown to them in the form of a screenshot. We employed a between-subject 2×1 experimental design with one control group, who saw the regular Instagram prompt asking for access to their address book (see Figure 3-3, left side), and one experimental group, who saw the same prompt enriched with an interdependent privacy salience nudge (IPN, right side).

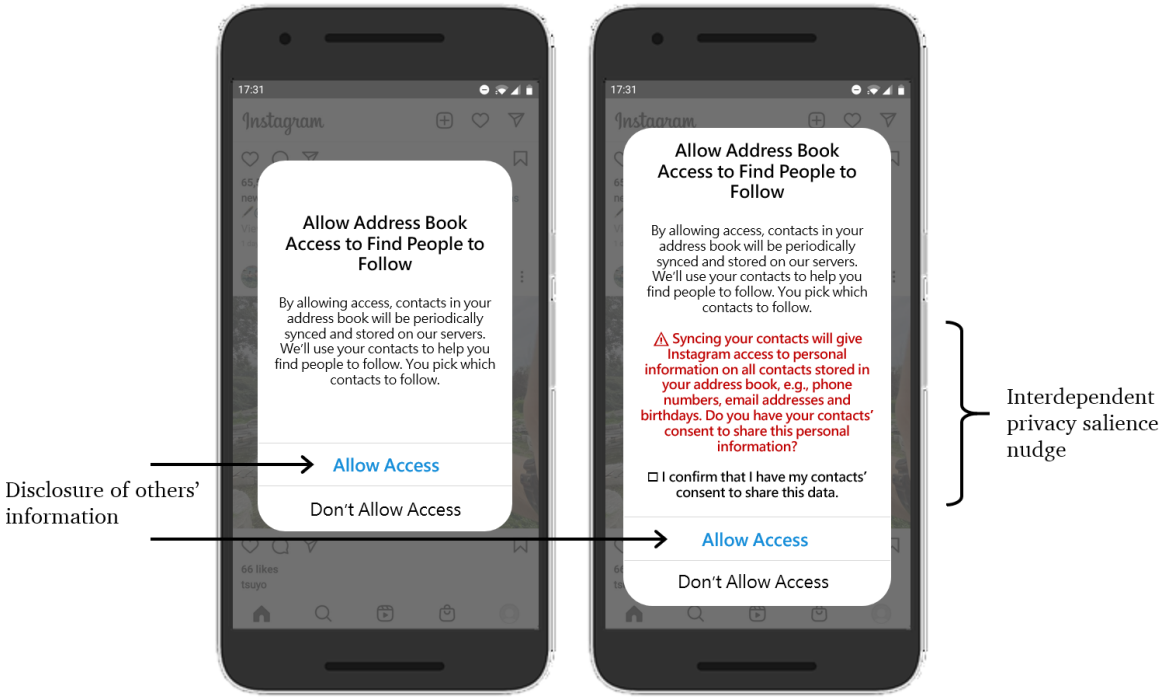


Figure 3-3. Mobile screenshots of Instagram prompt. Control group (left) and treatment group with interdependent privacy salience nudge (right)

Drawing on prior literature on privacy nudging, we designed our interdependent privacy salience nudge (IPN) with the following ideas in mind: Building on Kamleitner and Mitchell (2019)'s suggestions for interventions to improve interdependent privacy protection, we aimed to implement additional steps of decision control into the data transfer process by including an opt-in mechanism that requires users to actively confirm that they have their contacts' consent to share their personal information. Furthermore, we provide examples of the actual data about to be transferred to Instagram by specifying that it includes the phone numbers, email addresses and birthdays of all contacts stored in one's address book, to increase both the salience of the data and the data transfer as well as the salience of the other individuals involved. We hence designed our IPN with the aim to increase both users' RE and their RC. Our nudge design is in alignment with prior research on users' preferences regarding the design of privacy nudges by employing a default mechanism (i.e., allowing access is by default not possible without opting

in) along with color (red font), framing (warning sign), and privacy-related information (Schöbel et al., 2020). These design choices are also in agreement with the design principles for effective privacy nudging provided by (Barev et al., 2020).

In both the treatment and the control group, participants were asked how they would like to proceed with the Instagram prompt. After having selected one of two options, they were asked to shortly state why they chose the respective option in a free-text field. They were then redirected to our post-experimental questionnaire, where we recorded our mediation constructs, control variables, and socio-demographic information. Lastly, participants had the option to give feedback on the experiment on a voluntary basis, were debriefed, and informed that they finished the study.

3.4.2 Measured Variables

Measurement of the dependent variable

In our experiment, participants chose between the following two options on how to proceed with the Instagram prompt: (1) "Press 'Allow Access' to sync my contacts" for the control group and "Check the box 'I confirm that I have my contacts' consent to share this data' and press 'Allow Access' to sync my contacts" for the treatment group, respectively; (2) "Press 'Don't Allow Access' and not sync my contacts" for both groups. We measured our dependent variable, that is, participants' disclosure of others' information (DOI) by capturing if they chose to "Allow Access" (which we counted as "1") or "Don't Allow Access" (which we counted as "0").

Scale development for RE, RC, and RS

To measure our three mediation constructs RE, RC, and RS, we developed a measurement instrument based on the 3R Interdependent Privacy Protection Framework. In line with previous literature on scale development (e.g., MacKenzie et al., 2011; Moore and Benbasat, 1991), we started the process with a conceptual definition of the constructs, which was provided in detail by Kamleitner and Mitchell (2019). We then created a list of eight candidate items per construct that we thought to be suitable to represent the respective construct. Next, we asked six experienced researchers of the field of information systems to sort our items into the three constructs (RE, RC, RS), and to give feedback on the understandability of each item. This allowed us to assess the content validity of the items, to confirm the clustering into constructs, and to refine our wording. In a pretest experiment with 50 participants, we then evaluated the reliability of our items using Cronbach's Alpha (Cronbach, 1951). Based on the pretest results,

we again refined our measurement instrument and finally chose 4 items per construct to use in our experiment (see Table 3-4 in the Appendix).

Control variables

In addition to the constructs presented in our research model, we measured several alternative drivers of users' disclosure of others' information as controls in our experiment. Drawing on previous literature on users' information disclosure (e.g., Dinev and Hart, 2006; Krasnova et al., 2012), we measured participants' general privacy concerns (Pavlou et al., 2007; Smith et al., 1996) towards Instagram. Furthermore, we collected information on subjects' Instagram use, as well as sociodemographic information (i.e., gender, age, education, and nationality). Lastly, we measured users' normative beliefs towards disclosing others' information online (Primack et al., 2008). For a full list of all items used in our questionnaire, please refer to Table 3-4 in the Appendix.

3.4.3 Data Collection and Sample

In line with previous research, we recruited 349 participants via Prolific, a platform for recruiting subjects for social and economic science experiments (Palan and Schitter, 2018). All participants were EU citizens and were pre-screened as Instagram users by Prolific. Subjects were payed $0.53 (USD) for their participation. We excluded 19 participants because they failed to answer our attention check correctly (11 participants) or finished the study in less than half of the average completion time (8 participants), resulting in our final sample of 330 participants. Of the subjects in our study, 174 identified as women, 154 as males, and 2 as other. Participants exhibited an average age of 29.4 years, with 57% being younger than 25 years and 4% being older than 44 years. Our sample included 20 nationalities of the EU, with 95% of participants stating that they used Instagram at least several times a week, and 82% using it every day.

3.4.4 Serial Mediation Analysis

In our data analysis, we employ a serial mediation model with our three mediators RE, RC, and RS. In contrast to parallel mediation, where two or more mediators are hypothesized to explain the effect of an independent variable on a dependent variable while the mediators themselves do not causally influence one another, serial mediation describes two or more mediators that are linked together in a causal chain (Hayes, 2018). In our research model, we investigate the direct and indirect effects of our IPN on users' DOI while modeling a process in which the IPN increases RE and RC (the latter both directly and indirectly through RE), RC in turn feeds into

RS, concluding with DOI as the final consequence. We hence empirically test Kamleitner and Mitchell (2019)'s 3R framework, who have postulated that RE is causally prior to RC, which is causally prior to RS. The serial mediation approach is most fitting to explore research contexts where temporally ordered stages are central to theorizing (e.g., Casciano and Massey, 2012; Valentine et al., 2014).

3.5 Results

3.5.1 Measurement Validation

To confirm the random assignment of participants across our two experimental conditions (IPN absent vs. present), we performed a series of one-way ANOVAs for all control variables. There were no significant differences in gender ($F = 0.005$; $p > .1$), age ($F = 1.82$; $p > .1$), education ($F = 1.70$; $p > .1$), nationality ($F = 1.59$; $p > .05$), privacy concerns ($F = 0.94$; $p > .1$), normative beliefs ($F = 1.22$; $p > .1$) or Instagram use ($F = 0.32$; $p > .1$) among the two experimental conditions. It is therefore reasonable to conclude that the random assignment of participants to our conditions was successful, and that the participants' demographics and relevant controls did not confound the effects of our experimental manipulations. We assessed our item scales for reliability using Cronbach's alpha (Cronbach, 1951), which yielded values greater than 0.88 for all constructs (see Table 3-4 in the Appendix).

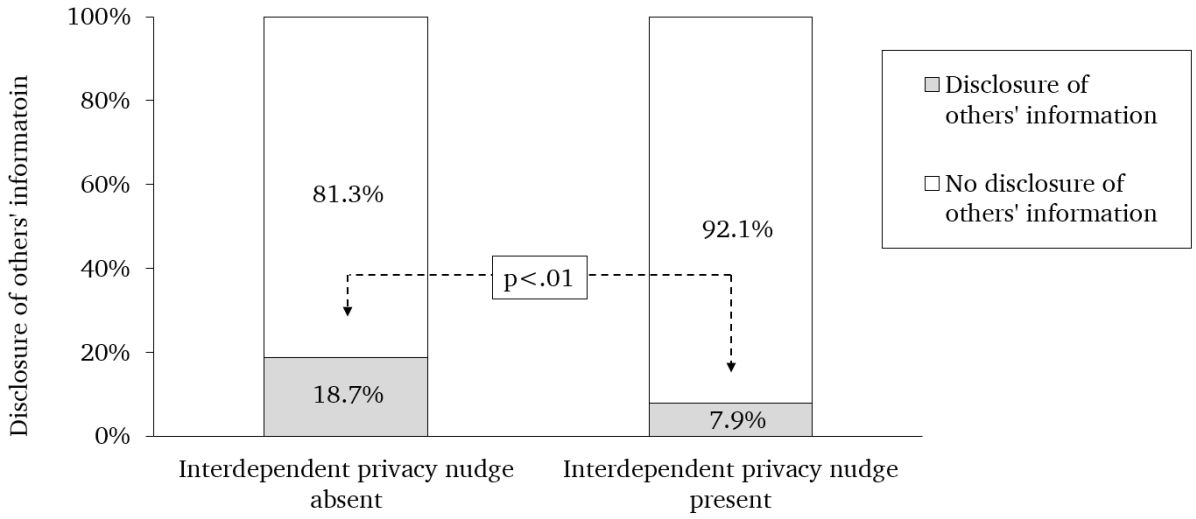


Figure 3-4. Disclosure of others' information across experimental groups; N = 330

3.5.2 Direct Effect of the IPN on Users' Disclosure of Others' Information

Of the 330 participants, 13.3% chose to disclose others' information by synchronizing their contacts with Instagram. A significant two-proportion Z test revealed that participants' choice varied across the two experimental conditions ($\chi 2 = 8.25$; $p < .01$), see 3-4: In the control group (IPN absent), 18.7% of participants chose to disclose others' information, whereas among participants who received the treatment (IPN present), only 7.9% chose to do so. In order to test our hypothesis H1, we conducted a binary logistic regression on our dependent variable DOI without and with control variables (Table 3-1).

	Binary logistic regression without controls			**Binary logistic regression with controls**		
Construct	B	*SE*	Exp(B)	B	*SE*	Exp(B)
Intercept	–1.47***	.20	.23	–.74	1.25	2.10
Manipulation						
IPN	–.98**	.35	.38	–.98*	.38	.38
Controls						
Privacy concerns	-	-	-	–.68***	.13	.51
Gender (male)	-	-	-	1.03**	.37	2.80
Age	-	-	-	–.46	.255	.63
Education	-	-	-	.30	.16	1.35
Nationality	-	-	-	–.03	.03	.97
Instagram use	-	-	-	–.32	.38	.73
Model fit						
Log Likelihood	–125.35	-	-	–103.48	-	-
Nagelkerke R^2	.05	-	-	.27	-	-
Omnibus χ^2	8.47**	-	-	52.20***	-	-
Note: * $p < .05$; ** $p < .01$; *** $p < .001$; N = 330						

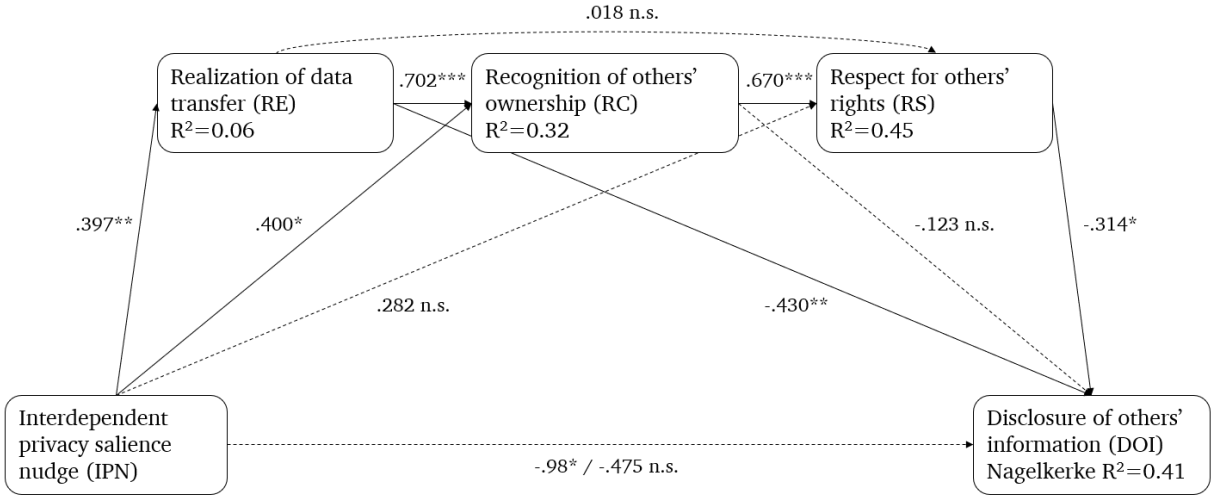
Table 3-1. Logistic regression analysis on participants' disclosure of others' information (DOI)

We examined the main effects of the IPN and any potential effect of the controls on participants' DOI. The results of our regression analysis demonstrate a significant negative effect ($B = -0.98$; $p < .05$; $Exp(B) = 0.38$) of the IPN on participants' DOI. Participants that were confronted with the IPN were hence 62% less likely to disclose others' information than when the nudge was absent, **supporting H1**.

3.5.3 Serial Mediation Analysis Through the Lens of the 3R Framework

Our remaining hypotheses focus on the explanatory mechanism through which the formation of users' decision to DOI takes place. In H2a and H2b, we hypothesized that the introduction of an IPN will increase users' RE as well as RC. In H3, we then argued that the IPN influences users' DOI negatively through a three-stage mediation via RE, RC and subsequently RS. To evaluate our hypotheses, we performed a serial multiple mediation analysis using Hayes (2018)'s PROCESS macro (version 3.5). We applied model 6 and entered RE, RC and RS as potential mediators between the independent and dependent variable while controlling for all direct and indirect paths. Additionally, we controlled the dependent variable as well as all mediators for participants' socio-demographics. Furthermore, we controlled the dependent variable as well as RE and RC for participants' privacy concerns, and, drawing on Kamleitner and Mitchell (2019), RS for participants' normative beliefs. We estimated our model using a bootstrapping approach based on 10,000 samples and 95% bias-corrected confidence intervals for the indirect effects. Figure 3-5 illustrates all direct effects as well as the explained variance of each constructs in our model. For a detailed stepwise presentation of all mediation effects, please refer to Table 3-5 in the Appendix. The model revealed a positive and statistically significant direct effect of the IPN on participants' RE ($B = 3.97$; $p < .01$). Furthermore, we found a positive and statistically significant direct effect of the IPN on participants' RC ($B = 0.400$; $p < .05$). This **corroborates our hypotheses H2a and H2b**, and implicitly confirms that our experimental treatment worked as intended.

In addition, Table 3-2 sheds further light on the indirect effects of RE, RC and RS on participants' DOI. We found evidence of three significant mediation paths, indicated by estimates of effect sizes that did not include zero in the given confidence interval. Path (1) demonstrates a significant indirect effect of the IPN on DOI through RE alone (effect size = –0.170; CI = [–0.4325, –0.0119], which has not been theorized in our research model. Path (6) consists of two significant specific indirect effects, namely through RC and RS as mediators (effect size = –0.084; CI = [–0.2247, –0.0004]. Furthermore, the direct effect of RC on DOI in our model is statistically insignificant ($p > .05$), suggesting a complete mediation through RS. Lastly, path (7) reveals a significant indirect effect of all three mediators RE, RC and RS (effect size = –0.059; CI = [–0.1619, –0.0023], hence **supporting our hypothesis H3**. As the direct effect of IPN on DOI ($B = –0.98$; $p < .01$) became insignificant after entering RE, RC and RS as mediators ($B = 0.475$; $p > .1$), this represents a full mediation through the 3R mechanisms (Hayes, 2018).



Note: The first coefficient on a given path represents the direct effect without the mediators in the model; the second represents the direct effect when the mediators are included in the model.

*** p < .001; ** p < .01; * p < .05; n.s. not significant; N =330

Figure 3-5. Results from the serial multiple mediation analysis

Indirect Effect Paths	**Effect size**	**Boot SE**	**BootLLCI**	**BootULCI**
(1) IPN → RE → DOI	**-.170**	**.107**	**-.4325**	**-.0119**
(2) IPN → RC → DOI	-.049	.086	-.2608	.0848
(3) IPN → RS → DOI	-.089	.076	-.2724	.0155
(4) IPN → RE → RC → DOI	-.034	.056	-.1592	.0659
(5) IPN → RE → RS → DOI	-.002	.016	-.0350	.0330
(6) IPN → RC → RS → DOI	**-.084**	**.060**	**-.2247**	**-.0004**
(7) IPN → RE →RC → RS → DOI	**-.059**	**.041**	**-.1619**	**-.0023**
Note: Coefficients were computed based on serial multiple mediation analysis including all controls and using bootstrapping with 10,000 samples and a 95% bias-corrected confidence interval (LLCI = Lower Limit / ULCI = Upper Limit of Confidence Interval) (Hayes, 2018). Significant indirect effects are marked in bold.				

Table 3-2. Results from the serial multiple mediation analysis, indirect effects

3.5.4 Post-hoc Analysis of Qualitative Results

After the participants of our experiment had decided how they would like to proceed with the Instagram prompt (by choosing either "Allow Access" or "Don't Allow Access"), we asked them to shortly state why they decided the way they did. This provided us with qualitative free-text answers and hence richer insights into participants' motives. Drawing on literature on the analysis of qualitative data (Mayring, 2014), we chose an inductive approach and coded all

answers into categories that illustrated a reflection of the data material. Some statements were sorted into more than one category, and some statements were too imprecise and therefore not sorted into any category. Table 3-3 gives an overview of participants' motives with exemplary statements.

Participants who disclosed others' personal information ("Allow Access")			
Code	**Number of statements**		**Exemplary statements**
	IPN **absent** (N=31)	IPN **present** (N=13)	
Gain more followers without much effort	21	9	*"It would be easier to find people that I know."* *"Because it's gonna help me add more contacts."*
Trust in Instagram	5	1	*"I think Instagram is safe enough to be trusted."*
Participants who did not disclose others' personal information ("Don't Allow Access")			
Code	**Number of statements**		**Exemplary statements**
	IPN **absent** (N=135)	IPN **present** (N= 151)	
Own privacy concerns towards Instagram	71	75	*"I don't like sharing my personal information to large companies."* *"I don't know how safe my data is with Facebook."* *"I try to avoid anything that barges into my privacy (more than social media already does)."*
Feature does not create added value	44	54	*"[…] I don't want to sync my contacts because I don't really care for that."* *"I don't have any interest in following people from my contact list."*
Respect for contacts' privacy rights	8	33	*"It does not seem right to share other peoples' data without their consent."* *"I have no right (or intend) to share personal data of my contacts with a company that makes money out of it."* *"Because the data in question does not belong to me."* *"I don't want to compromise my contacts' private info."*
Privacy protection towards contacts	15	15	*"I don't want every contact on my phone to see or find my Instagram account."* *"[…] I'd rather keep my Instagram private and prefer to not be able to be found by everyone in my contact list."*

Table 3-3. Analysis of participants' free-text statements

As for participants who disclosed others' personal information by allowing Instagram access to their contacts, the most stated reason was the wish to gain more followers without much effort. For example, one participant stated: "Because it helps to find people from my contacts without having to ask them for their username." Others seemed to have potential privacy concerns in mind, but found Instagram to be "safe enough to be trusted", or the prompt to be "relatively trustworthy".

Participants who chose to not allow Instagram access to their contacts stated several motives not to do so. While around one third stated that they "just don't need that feature" or have "no interest in it", privacy protection seemed to play a role in several ways. On the one hand, 146 participants stated that they had concerns regarding their own privacy towards Instagram (e.g., "I don't know how safe my privacy is with Facebook", or "I value my privacy more than having followers"). On the other hand, several participants expressed the wish to protect their privacy towards their contacts by stating, for example, that they "don't want every contact on [their] phone to see or find [their] Instagram account", or that they "rather keep [their] Instagram private". Lastly, and most interesting to our research, 41 participants (control group: 8, treatment group: 33) stated that they chose to not allow Instagram access to their address book due to their respect for their contacts' privacy rights. Participants' motives reflected our construct RS, for example, "I have no right […] to share personal data of my contacts […]" or "I don't want to compromise my contacts' private info", which at the same time implicates their RE and RC. The qualitative results hence further confirm our theoretical model.

3.6 Discussion

The disclosure of our own personal information through others increasingly threatens our privacy, specifically in the context of online platforms. Leading privacy researchers have hence called for a more versatile, multilevel understanding of privacy that acknowledges the complexity of sophisticated digital environments in order to be able to explore concepts such as, for example, interdependent privacy (e.g., Bélanger and James, 2020; Cao et al., 2018; Lowry et al., 2017). Whereas prior works on interdependent privacy have yielded important insights into the phenomenon of interdependent privacy, what has been missing is a deeper understanding of why and how users decide to either protect or violate others' privacy when interacting with online platforms, as well as an investigation of effective remedies. Our study revealed that an additional step of decision control in the form of an interdependent privacy salience nudge can decrease the likelihood that a user discloses others' information by 62%. Moreover, our results indicate a serial mediation through users' RE, RC, and RS when forming

the decision to disclose others' personal data. Our study holds important implications for the emerging field of research on interdependent privacy in both theory and practice, which we will elaborate on in the following.

3.6.1 Contributions to Interdependent Privacy Literature

We believe that our study contributes to interdependent privacy research in three particular ways. Our first contribution lies in the investigation of the interdependent privacy nudge itself. To our knowledge, this is one of the first works to explore the efficacy of an interdependent privacy salience nudge in a user study. Our interdependent privacy salience nudge was designed to increase the salience of the data and the data transfer as well as the salience of potential co-owners of the data. To do so, we implemented an opt-in mechanism, framing, as well as transparent privacy-related information as nudging mechanisms in alignment with prior literature (Kamleitner and Mitchell, 2019; Schöbel et al., 2020). Our experimental results show that the implementation of our nudge yielded a 62% decrease in participants' disclosure of others' personal information. This suggests a need for more transparent and salient communication of interdependent privacy implications on online platforms.

A second, broader contribution of this study relates to the theoretical mechanisms through which our interdependent privacy salience nudge decreases individuals' disclosure behavior. Manipulating the salience of the data transfer and the salience of the other, our data indicates that users' protection of others' privacy rights unfolds through a complete serial mediation of three consecutive stages as theorized by Kamleitner and Mitchell (2019). Our results hence empirically validate the 3R framework. However, while Kamleitner and Mitchell (2019) have proposed RE, RC, and RS to be three consecutive and hierarchically dependent steps, we found that the interdependent privacy salience nudge directly increases both RE and RC, and that there is a significant mediation path IPN → RC → RS → DOI. These findings deviate from the 3R framework, in that RC can be increased without relying on RE as a prerequisite. Our results advance our understanding of how interventions, such as our IPN, can act as effective remedies against interdependent privacy violations.

Our third contribution is of methodological nature and lies in the development and validation of a measurement instrument to capture users' RE, RC, and RS when disclosing others' information. Our measurement instrument allows to zoom into the micro-level processes of users' decision-making and might be useful for future studies in this field.

3.6.2 Implications for Practice

Our findings suggest implications for both policy-makers and online platform user interface designers. Current regulations, such as the European Union GDPR, do not sufficiently consider interdependent privacy infringements (Kamleitner and Mitchell, 2019; Symeonidis et al., 2018). This presents a loophole for online platform providers to intrude individuals' privacy rights through their peers. While online platforms ask the user for their consent (e.g., "Allow Access") when sharing others' information, the numerous data subjects that are involved in the data transfer are neither notified nor given the possibility to opt out. As demonstrated in our vignette scenario, around one fifth of Instagram users would give Instagram access to their address book in a real-world scenario, hence disclosing potentially hundreds of names, phone numbers, email addresses, and the like. Our results reveal useful insights for regulators, as they show the potential of providing users with design elements that increase the salience of the data transfer and the salience of the other. While, in a privacy-wise ideal world, users would not be allowed to share others' personal data without each of their co-owners' consent, this is not feasible in today's interconnected environment. However, we show that enabling users to make an informed decision about the disclosure of others' personal information can reduce interdependent privacy violations significantly. The significant indirect effect path (IPN → RE → DOI) reveals that social online platforms have potential for improvement in transparently communicating their practices for data collection and usage in general, since users seem to not be aware of the fact that data as a good is being transferred to the platform when, e.g., synchronizing contacts, let alone that this data belongs to other individuals. The introduction of the GDPR, which demands for mandatory opt-in mechanisms when giving access to one's own information, has proven to improve the transparency and visual representation of organizations' privacy policies (Linden et al., 2020). We argue that a future refinement of the GDPR should include mandatory and informative opt-in mechanisms when disclosing others' data, and hope that our work can inform future policy-making.

As for online platform providers, prior research has demonstrated that users' privacy concerns impact their choice of and behavior on online platforms (Gal-Or et al., 2018; Liu et al., 2021; Tsai et al., 2011) to the point where businesses might be able to leverage privacy protection as a selling point. The evaluation of our interdependent privacy salience nudge can serve as a starting point for the design of user interfaces where users can make informed decisions regarding interdependent privacy.

3.6.3 Limitations and Directions for Future Research

Despite the aforementioned theoretical and practical contributions of this research, our study is subject to several limitations, which open up a series of exciting venues for future research. Whereas our interdependent privacy salience nudge can serve as a starting point, we acknowledge that a more detailed evaluation of the individual nudging mechanisms (e.g., the default mechanism, the warning sign, or privacy-related information) is vital for the design of real-world interdependent privacy interventions. Prior research revealed that user-oriented information security and privacy interventions face issues such as information overload and habituation (Reeder et al., 2018; Vance et al., 2018), which have to be carefully navigated when aiming to support both individual and interdependent privacy decisions.

Our second limitation is of methodological nature. It has been argued that a research design where mediators are measured (such as the ones chosen in our study) as opposed to manipulated, is problematic to justify causality (Pirlott and MacKinnon, 2016). Since participants self-select to levels of the mediating variables, they are not randomly distributed across mediator levels, and the relationship between the mediators and the dependent variable can be correlational. Our ability to infer that our three mediators indeed caused DOI is hence limited, and our measurement of DOI might be subject to alternative explanations.

Furthermore, scholars might wish to extend our study by assessing the formation of our mediating variable RS in more detail. Kamleitner and Mitchell (2019) have suggested social norms and self-interest as important forces influencing RS, which has not been covered in our study. RS might also be subject to cultural factors: We conducted our study drawing on a sample of EU citizens, hence in countries with relatively similar cultural backgrounds. However, prior research has demonstrated the intercultural dynamics of privacy calculus on social networking sites (Krasnova et al., 2012). Accordingly, we encourage future research to further explore interdependent privacy protection across cultures.

Lastly, even though we designed our vignette experiment with the goal to represent a realistic scenario by employing an actual Instagram prompt as well as a sample of real-world Instagram users, our study relied on hypothetical and cross-sectional observations, and hence did not allow us to investigate users while they were actually deciding on interdependent privacy protection. To further strengthen the validity of our findings, we invite future research to apply complementary research methods, such as randomized field experiments.

3.7 Conclusions

This work investigates the formation of users' interdependent privacy decisions, and how such decisions can be supported by design elements such as an interdependent privacy salience nudge. We empirically validate Kamleitner and Mitchell (2019)'s 3R Framework of Interdependent Privacy Protection by revealing a serial mediation effect of our interdependent privacy salience nudge on users' disclosure of others' information through their realization of the data transfer (RE), recognition of other's ownership (RC) as well as respect for others' rights (RS). This answers our first research question, and sheds light on the steps that individuals have to take in order to be able to behave in an interdependent privacy-protecting manner. Addressing our second research question, we show that an interdependent privacy salience nudge employing an opt-in mechanism, framing, and transparent privacy-related information can support interdependent privacy protection on online platforms, with participants in the treatment group being 62% less likely to disclose others' personal information. This effect is reflected in users' qualitative statements, with participants expressing that "it does not seem right to share other peoples' data without their consent". Our study represents a starting point for future research on how to design usable and effective interventions for privacy protection in interdependent contexts.

3.8 Appendix

Construct	**Items**
	When deciding on how to proceed with the previous Instagram prompt, I was aware that syncing my contacts would imply…
Realization of data transfer (RE) (self-developed based on Kamleitner and Mitchell, 2019) (α = .875)	RE1: ...that I give Instagram access to data (such as names, email addresses or phone numbers). RE2: ...that I share data (such as names, email addresses or phone numbers) that I own with Instagram. RE3: ...that I transfer data (such as names, email addresses or phone numbers) from my belongings to Instagram. RE4: ...that I make data (such as names, email addresses or phone numbers) available to Instagram that they have not had before.
Recognition of others' ownership (RC) (self-developed based on Kamleitner and Mitchell, 2019) (α = .948)	RC1: ...that I give access to personal information of others. RC2: ...that I give access to data that has been shared with me by others. RC3: ...that I share data that belongs to others. RC4: ...that I share the information of others in addition to my own.
Respect for others' rights (RS)	RS1: ...that I treat others' privacy rights unfairly. RS2: ...that I disrespect others' privacy rights.

(self-developed based on Kamleitner and Mitchell, 2019) (α = .958)	RS3: ...that I treat others' personal data unlawfully. RS4: ...that, before sharing others' personal data, I should have obtained their consent.
Privacy concerns (Smith et al., 1996; Pavlou et al., 2007) (α = .889)	PC1: I am concerned about my privacy when browsing Instagram. PC2: I am concerned that Instagram is collecting too much information about me. PC3: It bothers me when Instagram asks me for personal information. PC4: My personal information could be misused when transacting with Instagram. PC5: My personal information could be accessed by unknown parties when transacting with Instagram. PC6: I have doubts as to how well my privacy is protected on Instagram.
Normative beliefs (Primack et al., 2008)	NB1: Among your peers, how socially acceptable is it to share others' information (e.g., names, phone numbers or email addresses) with platforms such as Instagram?
Note: 7-point Likert-type scales ranging from strongly disagree (1) to strongly agree (7) were used.	

Table 3-4. Measurement items

	M1 (RE)			**M2 (RC)**			**M3 (RS)**			**Y (DOI)**		
	b	SE	p	b	SE	p	b	SE	p	b	SE	p
X (IPN)	.40	.13	.003	.40	.16	.015	.28	.17	.107	–.48	.42	.258
M1 (RE)	-	-	-	.70	.07	.000	.02	.08	.829	–.43	.15	.003
M2 (RC)	-	-	-	-	-	-	.67	.06	.000	–.12	.14	.371
M3 (RS)	-	-	-	-	-	-	-	-	-	–.31	.13	.019
Constant	5.13	.48	.000	–.16	.67	.808	1.88	.67	.005	4.35	1.58	.006
	R^2=.06 F=2.98; p<.01			R^2=.32 F=18.98; p<.001			R^2=.45 F=28.52; p<.001			Nagelkerke R^2 =.41 Omnibus model $\chi 2$ = 83.02; p<.001		
All controls were included in the analysis.												

Table 3-5. Results from the serial multiple mediation analysis (coefficients and model summary information)

Chapter 4: Utilitarian versus Hedonic Motives in Cyber Threat Reporting

Title:	Why Do Employees Report Cyber Threats? Comparing Utilitarian and Hedonic Motivations to Use Incident Reporting Tools
Authors:	Anjuli Franz, Technische Universität Darmstadt, Germany
Published in:	International Conference on Information Systems (2022), December 9-14, Copenhagen, Denmark.

Abstract: Organizational cybersecurity is threatened by increasingly sophisticated cyberattacks. Early detection of such threats is paramount to ensure organizations' welfare. Particularly for advanced cyberattacks, such as spear phishing, human perception can complement or even outperform technical detection procedures. However, employees' usage of reporting tools is scarce. Whereas prior cybersecurity literature has limited its scope to utilitarian motives, we specifically take hedonic motives in the form of warm glow into account to provide a more nuanced understanding of cyber incident reporting behavior. Drawing on a vignette experiment, we test how the design features of report reasoning and risk indication impact users' reporting tool acceptance. The results of our mediation analysis offer important contributions to information systems literature by uncovering the dominant and under-investigated role of hedonic motives in employees' cyber incident reporting activities. From a practice perspective, our findings provide critical insights for the design of cyber incident reporting tools.

Keywords: Organizational cybersecurity, Cyber incident reporting, Hedonic motives, Warm glow

4.1 Introduction

Corporate cybersecurity issues challenge both research and practice since they are rooted in complex socio-technical systems, with human actors, technology, and processes acting as interconnected components (Zimmermann and Renaud, 2019). Prior information systems (IS) literature has predominantly labelled the human actor as the weakest link in the cybersecurity chain (e.g., Goel et al., 2017; Mitnick and Simon, 2003; Turel et al., 2021), that needs to be excluded, controlled, or trained in order to not present a hazard to organizational cybersecurity.

On the contrary, researchers have argued that this notion neglects the potential of human actors' capability to contribute actively to protecting and improving security (Zimmermann and Renaud, 2019; Kirlappos et al., 2013). Recent works have hence called for a paradigm shift from the human-as-a-problem to a human-as-a-solution cybersecurity mindset (Vielberth et al., 2021; Zimmermann and Renaud, 2019). No longer viewing the human "as a problem to control, but rather as a solution to harness" (Zimmermann and Renaud, 2019, p. 175) allows to fully tap humans' potential as a vital player in defending organizations against cyberattacks.

One of the most powerful capacities of humans in supporting organizational cybersecurity is the detection and reporting of suspicious activity, such as phishing attempts or anomalous behavior of software or hardware, which we refer to as cyber incident reporting (Heartfield and Loukas, 2018). The reporting of such incidents is paramount for organizations since it allows for early cyberthreat detection, which can critically reduce recovery cost and effort (Greene et al., 2018a). Since sophisticated cyberattacks often are not automatically detectable (Vielberth et al., 2021), human perception can act as an important source of contextual information, and has even been observed to be a superior security sensor and early warning system compared to technical procedures (Heartfield and Loukas, 2018). Over the last years, corporations have hence started to implement reporting tools, such as a phishing reporting button in email software, where employees can effortlessly report suspicious activities to the information security department. Employees' usage of such reporting functionalities, however, is scarce. While social engineering penetration tests have revealed that 78% of all employees never fall for a simulated phishing email and could hence potentially act as cyber incident reporters (Widup et al., 2018), only 7% actually report such a phishing attempt (NCATS, 2018).

While understanding what motivates employees to report cyberthreats is crucial for designing effective reporting mechanisms, IS research contributed little insight on this matter as of yet (Briggs et al., 2017; Vielberth et al., 2021). Literature on cyber incident reporting is scant, and first approaches have limited their scope to a utilitarian perspective (e.g., Kwak et al., 2020; Jensen et al., 2017a). We argue that this limitation does not account for the complex phenomenon of cyber incident reporting due to two main reasons: First, in the wider field of organizational cybersecurity, the research dialogue has steered towards the role of socio-emotional motivations (e.g., pride, or affective connection to colleagues) in employees' security behavior (e.g., Karjalainen et al., 2019; Renaud et al., 2021; Posey et al., 2014). Imagine, for example, the satisfying and proud emotion of feeling pleased with oneself after detecting and reporting a sophisticated malicious email. These insights have not been employed in cyber

incident reporting research as of yet. Second, cyberthreat reporting often takes place through technology, such as reporting tools implemented in email software. Prior works have found hedonic motives to play a crucial role in users' acceptance of technology (Van der Heijden, 2004; Wixom and Todd, 2005). However, to our knowledge, cyber incident reporting has not yet been studied through the lens of technology acceptance and its hedonic drivers.

We hence argue that, besides utilitarian motives, hedonic desires might play an important and hitherto under-investigated role in employees' reporting activities. From a practice perspective, shedding light on the underlying mechanisms of users' reporting behavior provides valuable insights for the design of cyber incident reporting tools striving to maximize employees' reporting activities. In this research work, we therefore intend to investigate the following two research questions:

1: How do utilitarian vs. hedonic factors influence employees' intention to use cyber incident reporting tools?

2: What are resulting implications for affordances that such reporting tools should offer?

To address these research questions, we conducted an online vignette study. Participants were presented with a self-developed email reporting tool equipped with two different design features, signaling the affordances of report reasoning (RR) (e.g., expounding one's reason to believe that the email is malicious) and risk indication (RI) (e.g., categorizing the report as a priority). The experiment was followed by a questionnaire, where the participants expressed their intention to use the email reporting tool. Furthermore, we measured the two constructs perceived usefulness (Davis, 1989) and warm glow of giving (Iweala et al., 2019; Andreoni, 1990) as mediators, representing participants' utilitarian and hedonic motives for using the reporting tool, respectively. Our results provide empirical evidence of the mechanism of both perceived usefulness and warm glow in affecting participants' intention to use an email reporting tool, with the hedonic feeling of warm glow contributing more strongly than perceived usefulness.

Our paper contributes to IS literature in general and cyber incident reporting research in particular: First, this research suggests that the concept of warm glow of giving might be a hitherto under-investigated IS continuance construct, which can play a pivotal role to enhance users' acceptance of otherwise utilitarian information systems. Second, this paper provides a novel perspective on organizational cybersecurity by challenging the prevalent assumption that purely utilitarian motives drive employees' intention to support their organization's security efforts (e.g., Hsu et al., 2015; Herath and Rao, 2009). By uncovering the dominating role of

employees' hedonic motives, we offer an important contribution to our understanding of why employees report cybersecurity incidents. Lastly, we shed light on affordances that foster both hedonic and utilitarian motives, and hence reveal important implications for the design of cyber incident reporting tools.

4.2 Theoretical Background

4.2.1 Behavioral Cybersecurity

Organizational cybersecurity is defined as the "efforts organizations take to protect and defend their information assets […] from threats internal and external to the organization" (Dalal et al., 2022, p. 5), and is distinguished by its interdisciplinary, socio-technical character (Zimmermann and Renaud, 2019; Craigen et al., 2014). While it is an organizational phenomenon, it heavily depends on the individual behavior of each employee, such as choosing secure passwords, locking one's computer screen when leaving one's desk, or not opening suspicious email attachments. Prior research has hence started to study behavioral cybersecurity, investigating, for example, the influence of psychological, social, emotional or cognitive factors on employees' protection of information systems' security (Dalal et al., 2022). On a cognitive level, employees' cybersecurity behavior has been explored through the lens of a rational cost-benefit analysis, studying the role of constructs such as users' perceptions of threat probability, response cost, rewards, or punishment severity in their security behavior (e.g., Herath and Rao, 2009; Hsu et al., 2015). By contrast, other research works have discussed users' affective needs as drivers of both compliance as well as noncompliance with information security policies (Karjalainen et al., 2019), or have investigated the role of socio-emotional factors such as ownership, involvement, fear, or personal pride in contributing to organizational security (e.g, Hsu et al., 2015; Posey et al., 2014). Whereas information security professionals seem to think more in terms of extrinsic motivations, such as punishments or rewards, as drivers for employees' security efforts, empirical data suggests that employees themselves are much more likely to be motivated by intrinsic motivations, such as organizational commitment, pride, or perceived responsibility towards their colleagues (Posey et al., 2014; Burda et al., 2020).

When regarding the role of the human factor in cybersecurity in general, previous IS literature has often considered the user to be the weakest link in the security chain, claiming that end-users lack security knowledge and awareness, are unmotivated to take responsibility, or simply lazy (Zimmermann and Renaud, 2019). Many research works have hence directed significant efforts to exploring, for example, how the human factor can be constrained and controlled via

information security policies (Cram et al., 2019; Li et al., 2021), how users' security knowledge and awareness can be increased via SETA programs (Bélanger et al., 2022; Silic and Lowry, 2020), or how user-centric design can support employees in engaging in secure behavior (Franz et al., 2021; Volkamer et al., 2017). Revealing intrinsic motives as a major driver for employees' information security efforts, however, opens the way for a new perspective on the human factor within the socio-technical cybersecurity system: The paradigm shift from "human-as-a-problem", who needs to be supported in preventing security incidents, to "human-as-a-solution", who can contribute actively to protecting the organization, allows organizations to fully reap human actors' capability to contribute to maintaining and enhancing cybersecurity (Zimmermann and Renaud, 2019). This is in accordance with Kirlappos et al. (2013)), who claim that the "comply or die" approach does not work for modern organizations, where employees collaborate and take initiative. In particular, several prior works have highlighted the capacity of human actors in reporting cyber incidents (Heartfield and Loukas, 2018; Vielberth et al., 2021), which is the topic of this study.

4.2.2 Cyber Incident Reporting

A cyber incident (or cybersecurity incident) is defined as an occurrence that misaligns the actual ownership and control rights of digital assets (which includes, for example, access, extraction, contribution, removal, or alienation) from the lawful ownership and control rights of these assets (Craigen et al., 2014). Cyber incident reporting describes a user's intentional report of a certain suspicion of, or relevant information about, such a cybersecurity incident, mostly via a computer-based reporting system (Vielberth et al., 2021). Early detection of such threats is paramount for organizations since it allows for fast incident response and containment, which can substantially reduce recovery cost and effort (Briggs et al., 2017; Greene et al., 2018a). Prior research has highlighted the capacities of human perception in complementing technical automated procedures (Heartfield and Loukas, 2018; Vielberth et al., 2021; Greene et al., 2018a). Particularly for social engineering attacks, that target the human factor via deception or masquerading techniques, human perception often outperforms technical filters: On the one hand, the vast majority of social engineering attackers leave little to no technical traces in their early stages and continuously evolve their attack patterns, exploiting, for example, zero-day vulnerabilities. This leaves technical heuristic detection capabilities with a meager starting basis, and a very limited view of potential threats through user interaction (Heartfield and Loukas, 2018; Vielberth et al., 2021). On the other hand, the detection of such attacks requires interpretation of both visual and behavioral information in their specific context, potentially

across multiple user-interface platforms (imagine, for example, a spear phishing email that contains a link to a cloud document). This makes human perception a more accurate security sensor than technical security systems, and hence an alluring candidate for actively contributing to cyberthreat detection (Heartfield and Loukas, 2018). While there will always be employees that fall for social engineering attacks and hence present a vulnerability for organizational cybersecurity, a single user who correctly detects and reports an incident can severely contribute to protecting the organization as a whole against cyberthreats.

Whereas organizations' cybersecurity can benefit profoundly from their employees' cyber incident reporting, employees' reporting activities are scarce (NCATS, 2018; Briggs et al., 2017). Prior works have hence called for research on the underlying motives that drive cyber incident reporting (Vielberth et al., 2021; Briggs et al., 2017). Empirical studies on this question, however, are scant. In the context of phishing reporting, Kwak et al. (2020)) have tackled the issue through the lens of Social Cognitive Theory, and have found that users' self-efficacy, cyber security self-monitoring, and expected negative outcomes influence their reporting motivation. Under the umbrella of theory from knowledge management and crowdsourcing, Jensen et al. (2017a)) have observed that public attribution and validation of successful phishing reports incentivizes employees to report their suspicions of malicious emails more frequently. Qualitative insights by Burda et al. (2020)) have suggested that reasons for reporting relate to the perceived sophistication of the attack, where users who assess themselves to have a higher sense of responsibility and threat awareness have expressed the motivation to safeguard less aware colleagues. These insights reflect the findings from the wider field of behavioral cybersecurity research, where both cognitive and affective factors have been observed to play a role in employees' security efforts (e.g., Karjalainen et al., 2019; Hsu et al., 2015; Herath and Rao, 2009).

4.2.3 Reporting Tools and Technology Affordances

From a tool perspective, the functionality to report suspicious or anomalous activity has found its way into most email software. This is in accordance with regulations such as ISO 27011, which requires the enablement of employees to report cyber incidents through suitable channels (e.g., A.16.1.2, ISO, 2013). Examining the reporting tool landscape in detail, however, reveals that little insights from research have found their way into practice as of yet. Most reporting tools are simple dialogue boxes with the two options to report an email as either spam or phishing, which then results in the email being forwarded to a predefined email address, and

the email being deleted from the user's account[8]. The current design hence likely does not acknowledge the underlying motives of employees' usage of such reporting tools, and arguably leaves much room for improvement. Affordance Theory (Gibson, 1979) offers a valuable means for user-centered analyses of technologies (Waizenegger et al., 2020; Piccoli, 2016; Tim et al., 2018). It relies on the assumption that individuals perceive their environment directly in terms of its potentials for action. Technology affordances are hence action possibilities afforded by a technology to its user (Gaver, 1991). If a technology application succeeds to offer salient affordances for users' psychological needs, this will typically motivate the use of such an application (Karahanna et al., 2018b). In this work, we test the effect of two reporting tool affordances on employees' usage intention.

4.2.4 User Acceptance and the Constructs of Perceived Usefulness and Warm Glow

Regarding user acceptance of technology in general, numerous research works have confirmed that both utilitarian and hedonic motives play a role in individuals' intention to use a certain technology (Van der Heijden, 2004; Dickinger et al., 2008). On the utilitarian side, the construct of perceived usefulness has been a central component of models for predicting user acceptance of technology for decades (Davis, 1989; Hu et al., 1999; Venkatesh et al., 2003). Its predictive ability on intentions to use technology has been supported by various research works in utilitarian contexts, and has often been employed as a counterpart to exploring hedonic motives for technology acceptance (e.g., Van der Heijden, 2004; Wakefield and Whitten, 2006). First introduced in the Technology Acceptance Model (TAM) by Davis (1985)), it describes the degree to which an individual believes that using a particular system will increase their job or task performance (Davis, 1989). Its theoretical grounding lies in the belief-intention relationships of the Theory of Reasoned Action (Fishbein and Ajzen, 1977), which suggests that users' beliefs influence their attitudes, which then lead to intentions, which in turn guide behaviors.

Hedonism refers to pleasure-seeking as motives for action (O'Shaughnessy and O'Shaughnessy, 2002). The core principle that distinguishes utilitarian systems from hedonic systems is that the first aim to provide only instrumental value to the user (e.g., enabling them

[8] For example, Lucy Security's PhishAlert plugin (https://wiki.lucysecurity.com/doku.php?id=phishing_incidents), KnowBew4's Phish Alert Button (https://support.knowbe4.com/hc/en-us/articles/360009629234-How-Do-I-Use-the-Phish-Alert-Button-for-Microsoft-365-), or ProofPoint's PhishAlarm (https://www.proofpoint.com/us/products/security-awareness-training/phishalarm-email-reporting).

to perform a certain task better), while the latter aim to offer a self-fulfilling value (e.g., experiencing fun or happiness when using the system) (Van der Heijden, 2004). Prior IS research has investigated hedonic constructs such as, for example, enjoyment (Van der Heijden, 2004; Dickinger et al., 2008), satisfaction (Wixom and Todd, 2005), or cognitive absorption (Agarwal and Karahanna, 2000) as predictors for technology acceptance. These hedonic constructs are driven by purely egoistic motives, that is, they provide users with enforcement of their own advantage without regard to others. In contrast, the hedonic construct of warm glow is based on altruistic behavior. The concept of "warm glow of giving" is based on Public Goods Theory (Andreoni, 1990) and reflects the satisfying "feeling people experience when performing an apparently altruistic act" (Iweala et al., 2019, p. 315). While an individual incentivized by pure altruistic motives is indifferent about the origin of the increased welfare of others, an individual driven by warm glow connects psychosocially with the recipient of the interaction, and receives a personal gain, such as a feeling of pride, enthusiasm, happiness, satisfaction, or boost of self-esteem, through the act of giving (Iweala et al., 2019; Gleasure and Feller, 2016). The concept of warm glow has mainly been limited to investigating charitable giving (Gleasure and Feller, 2016; Sutanto et al., 2021) and the influence of ethical claims on consumers' purchase intentions (Iweala et al., 2019; Lee and Charles, 2021), where warm glow givers have been described as "emotional altruists" (Singer and Ricard, 2015). Prior research on organizational cybersecurity has started to study the role of socio-emotional motivations, such as pride, in employees' security behavior (e.g., Karjalainen et al., 2019; Renaud et al., 2021; Posey et al., 2014). Investigating employees' cyber incident reporting behavior through the theoretical lens of warm glow might hence hold interesting insights for IS research.

4.3 Research Model and Hypothesis Development

Before presenting our research model, we primarily introduce two affordances of cyber incident reporting tools as potential design features to maximize user acceptance of such tools. The development of the two affordances investigated in this paper has been guided by both research and practice: Spear phishing incidents in our research department have sparked an extensive discussion among colleagues on what it feels like to detect a spear phishing attack in one's inbox, which has yielded results such as a feeling of surprise, excitement or satisfaction, as well as perceived superiority to others who might not be able to identify the email as phishing due to less security knowledge or context awareness. This notion is supported by prior literature, which has observed the role of involvement, ownership, or personal pride in cybersecurity-related behavior (Zimmermann and Renaud, 2019; Posey et al., 2014). When identifying a hard-

to-detect phishing attack, email users recognize the unique knowledge and valuable capabilities that they bring to this identification process, which neither technical controls nor IT experts might be able to contribute (Wash et al., 2021). Prior research, however, has not yet investigated these socio-emotional factors in the context of cyber incident reporting tools, neither have they been implemented by current phishing reporting tools (e.g., the examples from practice named earlier[8]).

When analyzing which exact psychological needs typically emerge from detecting a spear phishing attack, we agreed on (1) sharing details on the malicious email and how one successfully detected it with others (e.g., by showing colleagues a screenshot of the message, retelling the story of how one was almost tricked by criminals, or describing one's assessment of the specific attack characteristics), and (2) effectively warning others in case one thinks they might likely fall for the phishing attempt (e.g., by reporting the incident to the information security department or by telling colleagues directly about the incident). We then concluded that, out of these two psychological needs, current phishing reporting tools can only partly cater to the need of warning others (partly, since it remains unclear to the user if their report is handled with sufficient care and priority), and that the need to share one's own assessment of the attack characteristics remains largely unsatisfied. We hence developed two affordances, namely report reasoning (RR) and risk indication (RI), to address these needs. Report reasoning (RR) describes the possibility of explaining why one thinks that the reported occurrence is a cybersecurity incident. Regarding the reporting of a malicious email, for example, RR could be an affordance to explain which part of the email led to the assumption that it might present a security risk. Risk indication (RI) presents a way to indicate that the incident is high-risk, and that precautions should be taken immediately. Applied to the context of phishing, RI could be an affordance to flag a sophisticated attack, which might pose a severe threat to organizational cybersecurity, as a priority report.

To shed light on the effect of the reporting tool affordances RR and RI on our dependent variable intention to use, we propose a research model encompassing utilitarian and hedonic motives as drivers of employees' intention to use a cyber incident reporting tool. In the following, we expound upon each of the posited relationships as depicted in Figure 4-1.

On the left side of our model, we present RR and RI as independent variables. From a utilitarian perspective, employees who have detected a cyber incident will perceive the reporting of such an incident as a task they should fulfill in their role as a member of their organization. While RR allows users to pass on potentially important information (such as reporting a legitimately-

looking email in a suspicious context), RI enables them to ensure that others will be warned of a sophisticated attack before it spreads. The affordances to provide such relevant information on the incident through a reporting tool gives the tool an instrumental value, since users will feel like the tool helps them to perform the task of incident reporting better. This, in turn, will increase users' perceived usefulness of the reporting tool (Davis, 1989).

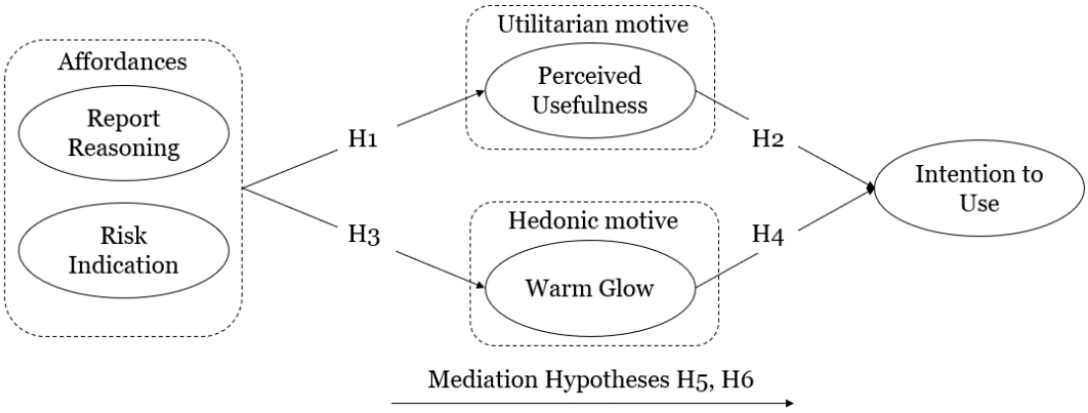


Figure 4-1. Research model

Furthermore, most employees do not possess expert knowledge on the identification of cyber incidents. We argue that RR and RI can provide guidance through one's own reflection of the security incident, and hence make the task of deciding whether or not to report an incident less difficult. Since prior research has identified users' perceived ease of use of a technology as an antecedent of perceived usefulness (Davis, 1989; Karahanna and Straub, 1999), we argue this mechanism reinforces the influence of RR and RI on perceived usefulness. Overall, we thus hypothesize that RR and RI have a positive effect on users' perceived usefulness of an incident reporting tool:

H1: The presence (vs. absence) of the affordances a) report reasoning, and b) risk indication is related to a higher level of perceived usefulness.

Perceived usefulness has in turn been confirmed to be a strong predictor of individuals' intention to use a technology (Davis, 1989; Venkatesh et al., 2003; Hu et al., 1999). Therefore,

H2: A higher level of perceived usefulness increases intention to use.

Beyond this cognitive rationale, prior research has observed the user-cybersecurity relationship to be driven by emotional and affective needs (Renaud et al., 2021; Karjalainen et al., 2019), such as the need to feel ownership of security decision processes (Hsu et al., 2015; Zimmermann and Renaud, 2019), or to feel validated when reporting a cyber incident (Jensen et al., 2017a). This holds especially for cybersecurity-aware employees, who tend to feel

responsible for safeguarding less aware colleagues (Burda et al., 2020). Both affordances RR and RI address these emotional needs. By enabling employees to interact directly with the information security department, RI and RR signal to users that their task expertise on a cyber incident is valued despite them not officially being security experts. Through RI, employees can take the role of a security advisor who can prompt the information security department to technically analyze a reported incident immediately. Being trusted with such security decisions invokes a feeling of active involvement in, and contribution to organizational welfare, which will enhance their perceived reputation. Furthermore, the affordance of RI will enhance employees' perception that their warning of others was effective, which will foster their satisfaction with the overall reporting process. Beyond that, RR addresses the urge to share one's story of the successful detection of a malicious threat as described at the beginning of this section. This can act as a way to indulge in the feeling of happiness and pride about one's achievement. Overall, we argue that RR and RI will act as a means to evoke and enhance feelings such as pride, satisfaction, happiness, and boost of self-esteem, which are an indication of the experience of warm glow (Iweala et al., 2019; Gleasure and Feller, 2016). We hence propose:

H3: The presence (vs. absence) of the affordances a) report reasoning, and b) risk indication are related to a higher level of warm glow.

Hedonic motives, such as satisfaction or enjoyment, have been identified as major drivers for usage intentions (Wixom and Todd, 2005; Van der Heijden, 2004), since they provide users with the self-fulfilling value of experiencing pleasure through technology usage. Building on IS literature, we hence argue that experiencing warm glow will motivate employees to report cyber incidents, which will result in a higher intention to use a reporting tool. We thus hypothesize:

H4: A higher level of warm glow increases intention to use.

In conclusion, we argue that the underlying motives of employees' cyber incident reporting are twofold. On the one hand, cyber incident reporting can be seen as a utilitarian act, where we assume employees to weigh their personal costs (e.g., spending time and effort) against benefits (e.g., increasing organizational cybersecurity). Since increasing the perceived usefulness of a reporting tool through RR and RI shifts the cost-benefit calculus in favor of the benefit, the presence of these affordances will result in a higher intention to use (Davis, 1989). On the other hand, employees' motives to report cyber incidents likely emerge from hedonic ambitions. Similar to a charitable donor giving towards a public good, an employee reporting a cyber

incident can experience a feeling of warm glow by psychosocially connecting with the recipient (that is, their organization or colleagues) of their altruistic behavior (Gleasure and Feller, 2016; Andreoni, 1990). RR and RI augment this psychosocial connection by feeling actively involved in contributing to organizational cybersecurity, which increases employees' feeling of pride, satisfaction, and happiness (Jensen et al., 2017a; Posey et al., 2014). The pursuit of the feeling of warm glow hence also drives their usage intention of a cyber incident reporting tool. As such, we suggest that both perceived usefulness and warm glow mediate the effect of our two affordances on intention to use:

H5: Perceived usefulness mediates the effect of the affordances a) report reasoning, and b) risk indication on intention to use.

H6: Warm glow mediates the effect of the affordances a) report reasoning, and b) risk indication on intention to use.

4.4 Methodology

With the goal to unravel the role of altruistic vs. hedonic motives in employees' intention to use a cyber incident reporting tool, we opted for an online vignette experiment in an email reporting context. We chose the vignette methodology since it permits to control for participants' personal experience and to avoid social desirability bias (Aguinis and Bradley, 2014), and because it has been validated as an effective technique for assessing users' perceptions of and reactions to cybersecurity-related conditions (Benlian et al., 2020; Warkentin et al., 2017). In our experiment, participants were asked to imagine they were employed at a fictional company called TradeFurnishings, which had experienced several cybersecurity issues through phishing or ransomware attacks in the past. Employees were hence asked to report unsolicited emails to the information security department using a reporting tool implemented in their email program. In our experiment, participants were then introduced into the functionalities of the current email reporting tool, which consisted of a report button in the menu bar of their email program, and a dialogue box with two radio buttons "report as spam" and "report as phishing". We decided to use this current tool as a baseline to avoid preconceived attitudes governed by participants' potential past interactions with real-world email reporting tools. In our study, participants were then informed that the information security department had implemented an updated version of the email reporting tool, and were presented with the novel functionalities. Here, we randomly assigned our sample to four conditions, yielding a 2x2 between-subject design.

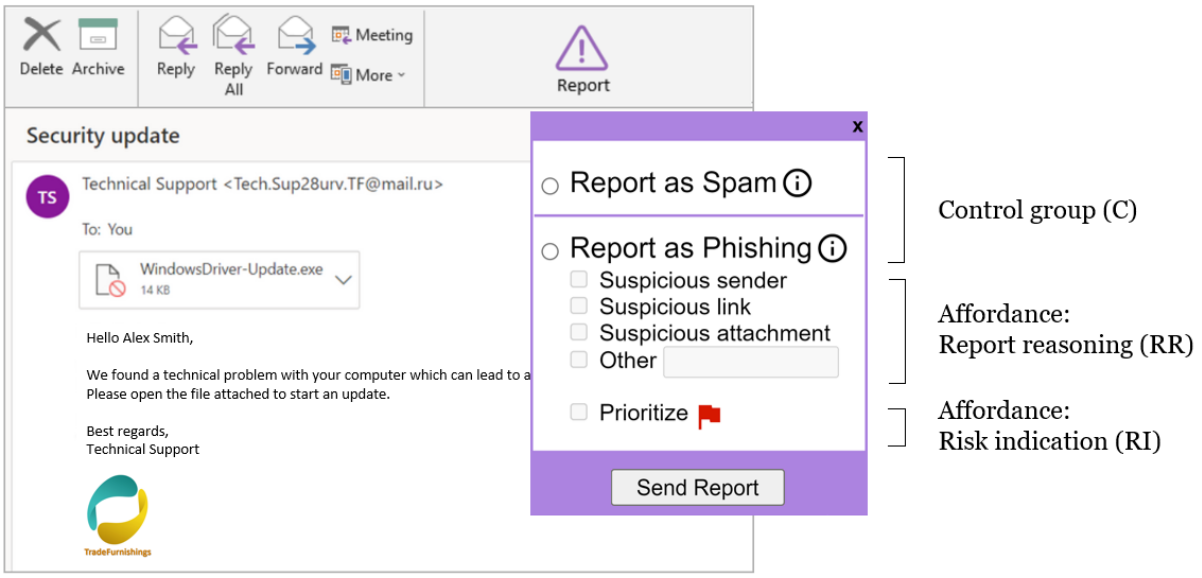


Figure 4-2. Exemplary Phishing Email with Email Reporting Tool Dialogue

For the control group C as well as all other groups, the previous tool was updated with an element where participants could access brief information on what is spam and what is phishing by hovering over an information icon. The treatment group RR was additionally given the opportunity to multi-select reasons why they thought that this particular email was malicious, e.g., because the sender or link seemed suspicious, hence reflecting the affordance of report reasoning. Participants were informed that their assessment helped the information security department to analyze the email. The treatment group RI could optionally flag the report with a priority tag, signaling the affordance of risk indication. Participants were told to use this priority tag if they believed that they were reporting a sophisticated malicious email that might pose a severe threat to their colleagues and organization, and that their vigilance enabled the information security department to take precautions immediately. Lastly, participants in group RR*RI were presented with a tool that included both affordances RR and RI, as depicted in Figure 4-2.

After familiarizing themselves with the updated email reporting tool, participants were presented with six consecutive emails, of which three were phishing emails, two were legitimate emails, and one was spam, and were asked to report them via the reporting tool if they perceived them to be phishing or spam. The three phishing emails ranged from mass to spear phishing, with background information from our vignette story (e.g., the TradeFurnishings logo or the name of the CEO) serving as masquerading techniques. The email depicted in Figure 4-2 was designed to be of medium difficulty to recognize as phishing.

Compared with the previous email reporting tool, how do you feel about the new email reporting tool? Please rank your agreement with the following statements.	
Perceived Usefulness (PU) (adapted from Davis, 1989) (α = 0.92)	PU1: The new email reporting tool enhances the effectiveness of employees' reports of unsolicited emails. PU2: I find the new email reporting tool more useful. PU3: The new email reporting tool addresses my organization's security-related needs better.
Warm Glow (GLO) (adapted from Iweala et al., 2019) (α = 0.96)	GLO1: Reporting emails with the new email reporting tool gives me a stronger pleasant feeling of personal satisfaction. GLO2: I am more satisfied with myself when I use the new email reporting tool. GLO3: Using the new email reporting tool, I feel happier contributing to TradeFurnishing's security. GLO4: I am more satisfied with myself when I make a contribution towards email security at TradeFurnishings.
Intention to Use (ITU) (adapted from Wixom and Todd, 2005; Taylor and Todd, 1995) (α = 0.94)	ITU1: I have higher intentions to use the new email reporting tool as a routine part of my job over the next year. ITU2: I plan to use the new email reporting tool more frequently. ITU3: I intend to use the new email reporting tool more often when receiving unwanted emails.
Note: All items were measured on a 7-point Likert scale, ranging from strongly disagree (1) to strongly agree (7). α represents Cronbach's Alpha (Cronbach, 1951).	

Table 4-1. Measurement items

Having processed the emails, participants completed a questionnaire on their perceptions of the email reporting tool. To operationalize our constructs, we used and adapted existing measures. The items for perceived usefulness, warm glow, and intention to use are presented in Table 4-1. Additionally, we measured demographics (gender, age) and control variables (affinity for technology, phishing identification expertise, average of emails received per day).

Our sample was drawn via Prolific, a crowdworking platform for recruiting subjects for scientific experiments (Palan and Schitter, 2018). All participants were pre-screened by Prolific as white-collar workers using technology at work more than once a day and speaking English fluently, and were payed US$0.82 for their participation. In total, 277 participants took part in our experiment. Responses from 43 participants who failed at least one of our attention checks were excluded, resulting in our final sample of 234 participants. The distribution across experimental groups is depicted in Table 4-2. Of the subjects in our study, 54.3% were women,

22.8% were between 25 and 34 years old, and 88.8% lived in the United States. To ensure that our participants were indeed randomly assigned to our four treatment groups, we conducted an ANOVA based on our sample demographics, which yielded no significant difference (all $p>0.05$).

The acceptance of our reporting tool within the experiment was high, participants largely reported phishing correctly at least once during the experiment (90.6%). To illustrate participants' interaction with the reporting tool, we employ the email depicted in Figure 4-2 as an example: The email was reported as phishing by 81.2% of all participants. Those participants who had the RR element available largely checked the box for suspicious attachment (80.6%), and partly for suspicious sender (53.8%). Of those participants who had the RI element available, 56.8% made use of the priority flag, indicating their assessment that the email is high-risk and should be analyzed by the information security department immediately.

Experimental groups	**C**	**RR**	**RI**	**RR*RI**
N	61	60	57	56
Correctly reported phishing at least once	87%	88%	95%	93%

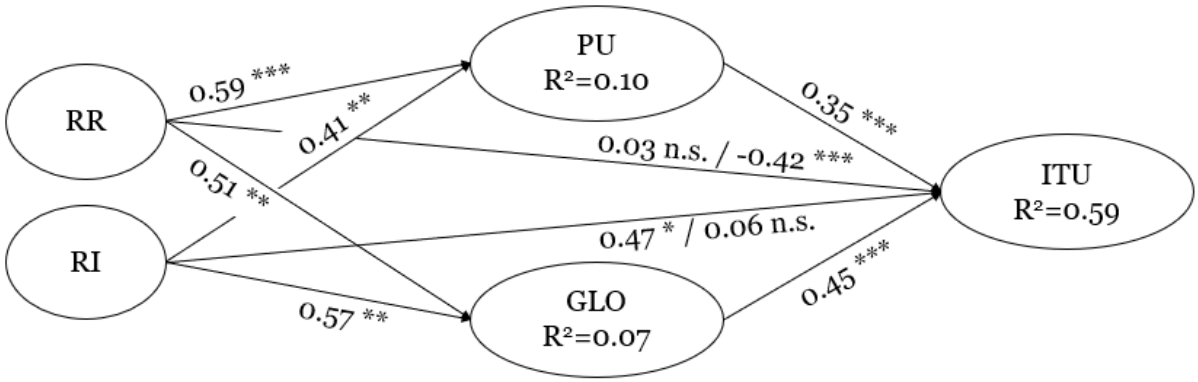
Table 4-2. Experimental groups

4.5 Results

To analyze our results, we first conducted a linear regression analysis with the presence vs. absence of our two affordances RR and RI as independent variables and intention to use as dependent variable, along with our control variables as covariates. The results indicate a positive direct effect of RI on intention to use ($\beta = 0.47$, $p < 0.05$). In contrast, we find no indication of a significant effect of RR ($\beta = 0.03$, $p > 0.05$), or the interaction term RR*RI ($\beta = -0.02$, $p > 0.05$), on intention to use. As for our control variables, our results suggest a positive effect of participants' affinity for technology ($\beta = 0.24$, $p < 0.001$) as well as gender ($\beta = 0.36$, $p < 0.05$; female = 1).

To test our hypotheses, we then entered perceived usefulness and warm glow as potential mediators in our model. Figure 4-3 shows the direct and indirect effects of our mediation model analysis. For perceived usefulness (PU), results of our regression model indicate a positive and significant effect of both RR ($\beta = 0.59$, $p < 0.001$) and RI ($\beta = 0.41$, $p < 0.01$). We therefore find **support for H1a and H1b.** The combined variance in perceived usefulness explained by the presence of our affordances RR and RI is 10%.

Furthermore, our analysis confirmed a positive and significant effect of both RR ($\beta = 0.51$, $p < 0.01$) and RI ($\beta = 0.57$, $p < 0.01$) on warm glow (GLO), thus **supporting H3a and H3b**. The regression model explains 7% of the variance in warm glow.



Indirect effects	Coefficient	SE	LLCI	ULCI
RR → PU → ITU	0.21	0.06	0.0886	0.3383
RR → GLO → ITU	0.23	0.08	0.0701	0.4101
RI → PU → ITU	0.14	0.06	0.0390	0.2710
RI → GLO → ITU	0.25	0.09	0.0866	0.4511

Note: N=234; ***p<0.001; **p<0.01; *p<0.05; n.s. not significant.
The first coefficient on a given path represents the direct effect without the mediators in the model; the second represents the direct effect when the mediators are included in the model. Coefficients for indirect effects were computed using bootstrapping with 10,000 samples and a 95% bias-corrected confidence interval. LLCI and ULCI denote the lower bound and upper bound of the confidence interval, respectively. All control variables were included in the analysis.

Figure 4-3. Direct and indirect effects of the mediation analysis

For the influence of perceived usefulness on intention to use (ITU), our results indicate a positive and significant effect ($\beta = 0.35$, $p < 0.001$), which is **in support of H2.** Moreover, warm glow has a significant positive influence on intention to use ($\beta = 0.45$, $p < 0.001$), hence **supporting H4**. Our final model explains 59% of the variance in intention to use.

Lastly, we conducted two mediation analyses using Hayes (2018))'s PROCESS macro (version 4.0), which is based on ordinary least squares regression. We provide results based on a bootstrapping approach with 10,000 samples and 95% bias-corrected confidence intervals for the indirect effects.

Our hypothesis H5 posited that the presence of RR and RI affects users' intention to use through perceived usefulness. Results of our mediation analysis reveal a positive indirect effect for both paths RR→PU→ITU (indirect effect = 0.21, CI = [0.09, 0.34]) and RI→PU→ITU (indirect effect = 0.14, CI = [0.04, 0.27]). As such, perceived usefulness mediates the effect of RR and RI on intention to use, **thus supporting H5a and H5b.**

H6 posited that the presence of RR and RI affects users' intention to use through warm glow. Our mediation analysis results indicate a positive indirect effect for both paths RR→GLO→ITU (effect size = 0.23, CI = [0.07, 0.41]) and RI→GLO→ITU (effect size = 0.25, CI = [0.09, 0.45]). Therefore, warm glow mediates the effect of RR and RI on intention to use, **thus supporting H6a and H6b.**

In summary, we find that the effect of RR and RI on intention to use can be explained by a parallel mediation through perceived usefulness and warm glow. Warm glow is likely a more dominant driver of reporting tool acceptance because the coefficient is higher in both the direct and indirect effects. The positive direct effect of RI on intention to use becomes insignificant when entering our two mediators into the model. This means that RI no longer affects intention to use when controlling for perceived usefulness and warm glow, which is often referred to as full mediation (Zhao et al., 2010). While our results indicated no significant direct effect of RR on intention to use, the direct effect becomes negative and significant ($\beta = -0.42$, $p < 0.001$) when entering perceived usefulness and warm glow into the model. This suggests a competitive mediation, and hence the existence of an omitted mediator that is competitive to the positive indirect effects of perceived usefulness and warm glow (Zhao et al., 2010).

4.6 Discussion

Organizational cybersecurity hinges on employees' security behavior. While employees have been considered a threat to cybersecurity for a long time (Zimmermann and Renaud, 2019), research has started to acknowledge their vast potential in cyber incident reporting (Heartfield and Loukas, 2018; Vielberth et al., 2021). Despite their potential, however, employees' reporting activities are scant, which leads to the assumption that current incident reporting tools do not fulfill employees' needs. Although prior works have recognized the importance of studying employees' acceptance of reporting tools, the underlying motives of cyber incident reporting have not yet been unraveled. While prior literature has limited its scope to utilitarian motives (e.g., Kwak et al., 2020), the main objective of our study was to specifically explore hedonic motives. Drawing on donation literature (Andreoni, 1990; Gleasure and Feller, 2016),

we employed the construct of warm glow to operationalize hedonic motives. Our research presents three important findings.

First, our investigation reveals both warm glow and perceived usefulness as key factors for employees' cyberthreat reporting behavior. The strong weight of warm glow (0.45) represents its critical role in reporting tool usage intentions, compared with perceived usefulness (0.35). Second, the results of our mediation analysis indicate that the two design features risk indication (RI) and report reasoning (RR) present a useful extension of current cyber incident reporting tools. For both features, we found significant positive indirect effects on employees' intention to use via perceived usefulness and warm glow. Lastly, our results suggest a competitive mediation for the effect of RR on intention to use. While our findings suggest a positive indirect effect through perceived usefulness and warm glow, the direct effect of RR on intention to use becomes negative when controlling for both mediators. This informs our theorizing of the possible existence of a omitted mediator with a negative sign in our research model (Zhao et al., 2010). While this can be pursued in future research, we speculate that potential candidates might be perceived effort or productivity loss: In comparison to RI, the feature of RR might be associated with higher effort by the user, since it requires more interaction. Conflicts with productivity have been found to be main reasons for non-compliance with security policies (Kirlappos et al., 2013; Sasse, 2015). Overall, these results provide a more nuanced understanding of cyberthreat reporting behavior and shed light on a vast potential for reporting tools to tap into.

4.6.1 Contributions to Theory and Practice

Our research offers two main contributions to the IS literature in general and to cybersecurity literature in particular.

First, this paper investigates the role of hedonic motivation in technology acceptance. While this has been extensively done by prior works, most authors have limited their scope to hedonic motives that are of rather egoistic nature, such as enjoyment (Van der Heijden, 2004), user satisfaction (Wixom and Todd, 2005), or cognitive absorption (Agarwal and Karahanna, 2000). These constructs describe experiences that provide users with an advantage without regard to others. Conversely, the concept of warm glow describes a hedonic experience based on altruistic behavior (Andreoni, 1990). To date, IS literature's interest in the role of warm glow has been limited to charitable behavior in purchasing or crowdfunding contexts (Gleasure and Feller, 2016; Lee et al., 2018). Drawing on our insights in this work, we argue that warm glow might hold interesting interactions embedded within technology in other research domains. We

hence call for research on this hitherto under-investigated IS continuance construct, which can have pivotal influence on users' acceptance of otherwise utilitarian information systems.

Second, this paper provides a new perspective on organizational cybersecurity. Prior IS literature has mostly assumed end-users to lack security knowledge, awareness, and motivation, thus presenting the weakest link in the security chain. While first works have started to acknowledge the power of the user in protecting organizational cybersecurity (Zimmermann and Renaud, 2019), most research has limited its scope to the prevailing assumption that purely utilitarian motives drive employees' intention to support their organization's security efforts (Hsu et al., 2015; Herath and Rao, 2009; Kwak et al., 2020). While utilitarian motives undoubtedly are a strong predictor of reporting tool usage, our empirical data uncovers the dominating role of hedonic motives in cyberthreat reporting behavior. This challenges the prevailing assumption of why employees report cyberthreats, and answers our first research question. With our findings, we additionally contribute to a more nuanced understanding of factors that explain employees' security behavior in general. Our drawing of the analogy between charitable behavior (Gleasure and Feller, 2016; Iweala et al., 2019) and organizational cybersecurity behavior can inform future theorizing.

Beyond our theoretical contributions, our paper provides important implications for designers of cyber incident reporting tools. Addressing our second research question, our analysis of the underlying motives of employees' reporting intentions uncovers that the design of cyber incident reporting tools should address both utilitarian and hedonic user needs. Informed design decisions can cater to both a strong feeling of perceived usefulness and an experience of warm glow in order to maximize continuance intention. While current reporting tools (such as the one in our experimental control group) do not foster users' hedonic needs, our two design features RR and RI provide a valuable example of how reporting tool design can harness the potential of employees' reporting capacities. While RI (that is, the option to flag reports as a priority) yielded a net positive effect on participants usage intentions and hence represents an attractive candidate for practice, RR apparently needs more finetuning. Furthermore, other mechanisms, such as bonuses or rewards, might be able to stimulate hedonic aspects of reporting cyber incidents.

4.6.2 Limitations and Future Research

We recognize limitations of our research, which hopefully provide opportunities for future works. First, we would like to highlight methodological limitations. Although we measured participants' behavioral intentions, our experimental setup did not allow to measure their actual

behavior. While previous studies have observed that the assessment of behavioral intentions provides a reasonable indication of their actual behavior (Venkatesh et al., 2012), we encourage future research to verify our findings through experiments in the field. Furthermore, methodological means such as manipulation checks for our two design affordances as well as the inclusion of further control variables such as perceived ease of use (Davis, 1989) would add to the robustness of our experimental data[9]. Second, our study was conducted in the context of email reporting. Although it is likely that our results are applicable to cyber incident reporting in other contexts, this limits the generalizability of our work. For example, our findings may not be applicable to cybersecurity incidents that require higher degrees of security expertise, or that employees are typically exposed to less frequently than to malicious emails. We therefore call for future research to replicate our findings in other cybersecurity contexts to confirm generalizability.

[9] We would like to thank the Associate Editor of this paper for these valuable suggestions.

Chapter 5: Behavior-based Measurement Instrument for Organizational Security Awareness

Title:	Spear Phishing 2.0: Wie automatisierte Angriffe Organisationen vor neue Herausforderungen stellen
Authors:	Anjuli Franz, Technische Universität Darmstadt, Germany Alexander Benlian, Technische Universität Darmstadt, Germany
Published in:	HMD Praxis der Wirtschaftsinformatik (2020), 57(3), 597-612.

Abstract: Vom ursprünglichen „Phishing = Passwort + Fishing" wandelt sich das Angriffsmuster durch neue Technologien zum boomenden Geschäftsmodell der cyberkriminellen Szene. Schadsoftware wie „Emotet" zeigt, dass automatisierte Spear Phishing-Angriffe Realität geworden sind und immense Schäden verursachen. Der Mitarbeiter rückt damit in den Fokus von IT-Sicherheitsmaßnahmen. Das Ziel dieses Beitrags ist es, einen Rundumblick zur aktuellen und zukünftigen Bedrohungslage durch Spear Phishing zu geben und konkrete Handlungsempfehlungen abzuleiten. Zur Messung der Security Awareness im organisatorischen Umfeld wird die Kennzahl „Employee Security Index" vorgestellt, welche das Sicherheitsbewusstsein von Mitarbeitern gegenüber Phishing-Angriffen standardisiert messbar macht. Es wurde ein Feldexperiment in einer deutschen Organisation durchgeführt, um die Verwundbarkeit der Belegschaft gegenüber Spear Phishing und die Wirksamkeit verschiedener Trainingsmaßnahmen zu untersuchen. Die erhobenen Daten werden mithilfe des „Employee Security Index" bewertet. Insgesamt verdeutlichen die Ergebnisse, dass neben technischen und organisatorischen Schutzmaßnahmen sowohl eine Schulung der Mitarbeiter als auch ein Umdenken nutzerverbundener Prozesse unabdingbar ist.

Keywords: Spear Phishing, Security Awareness, Social Engineering, Emotet, Faktor Mensch, Employee Security Index

5.1 Einleitung

Im Zeitalter der Digitalisierung stellen Phishing-Angriffe Unternehmen, Organisationen sowie Privatpersonen vor wachsende Herausforderungen (Benlian, 2020). Phishing ist ein Teilbereich von Social Engineering. Social Engineering bezeichnet Angriffsmuster, welche auf die Schwachstelle Mensch abzielen, um IT-Systeme anzugreifen. Cyber-Kriminelle geben sich

hierbei z.B. als vertrauenswürdige Quelle aus und nutzen E-Mails als Angriffsvektor, um Schadsoftware im Netzwerk zu platzieren, Zugangsdaten abzugreifen oder sich finanziell zu bereichern (Wright et al., 2014a).

Da Unternehmen zunehmend mehr in technische Schutzmaßnahmen investieren, ist der Weg über den „Faktor Mensch" für Angreifer oft der einfachere. Der Nutzer wird daher häufig als das schwächste Glied der IT-Sicherheitskette bezeichnet. Sogenannte Spear Phishing-Angriffe beschreiben dabei fortgeschrittene, zielgerichtete Phishing-Angriffe, welche individuell auf Personen oder Organisationen ausgerichtet sind. Spear Phishing war im Jahr 2019 der beliebteste Angriffsvektor bei Cyber-Angriffen (Symantec, 2019). Schaffen es die Angreifer, mit einer Spear Phishing-E-Mail Zugang zum Firmennetzwerk zu erhalten, geht der finanzielle Schaden schnell in die Höhe: Die durchschnittlichen Kosten eines solchen Vorfalls betragen für KMU etwa 1,4 Millionen Euro (Cloudmark, 2016).

Neben rein finanziellen Schäden sind bei Phishing-Angriffen häufig Produktionsausfälle, Reputationsschäden und Wirtschaftsspionage die Folge. Denkt man in Richtung Internet of Things (IoT), stellen hochvernetzte IT-Infrastrukturen ein äußerst lukratives Angriffsziel für beispielsweise DDoS (Distributed denial of service)-Angriffe dar (Hertel, 2017). Laut einer Studie des Digitalverbands Bitkom entsteht der deutschen Wirtschaft durch digitale Spionage, Sabotage und Datendiebstahl ein Schaden von 21 Mrd. Euro jährlich (Bitkom, 2018).

Dieser Beitrag gibt einen Einblick in die Thematik „Spear Phishing" und beleuchtet, welche Maßnahmen Unternehmen und Organisationen ergreifen sollten, um sich gegen aktuelle Bedrohungen zu schützen. Abschnitt 5.2 behandelt die Evolution von Phishing über die letzten Jahre und präsentiert Angriffsmuster realer Vorfälle aus 2019. Abschnitt 5.3 widmet sich dem Thema „Security Awareness", d.h. dem Sicherheitsbewusstsein von Mitarbeitern, und stellt verschiedene Trainingsansätze vor. In Abschnitt 5.4 präsentieren wir die Ergebnisse eines Feldexperiments, welches die Messung und Steigerung der Security Awareness gegenüber Spear Phishing in einer Organisation untersucht. Abschließend gibt Abschnitt 5.5 konkrete Handlungsempfehlungen für Informationssicherheitsverantwortliche.

5.2 Phishing: Eine kurze Evolutionsgeschichte

Von der ursprünglichen Definition „Phishing = Passwort + Fishing" wandelt sich das Angriffsmuster durch neue Technologien zum boomenden Geschäftsmodell der cyberkriminellen Szene. Angreifer nutzen automatisiert öffentlich verfügbare Informationen und setzen immer komplexere und glaubwürdigere Angriffsmuster ein. Dies fordert einen

gemeinsamen Kraftakt von technischen und organisatorischen Schutzmaßnahmen sowie aufmerksamen Mitarbeitern. Dieser Abschnitt gibt einen aktuellen Rundumblick zum Thema Phishing.

5.2.1 Besser, leichter, öfter: Wachsende Risiken durch Phishing

Im klassischen Sinn beschreibt Phishing das Abgreifen von Zugangsdaten auf gefälschten Login-Seiten, das heißt das „Fischen" von Passwörtern. Da die Angriffsmuster und -motive immer komplexer werden, versteht man heute den Begriff oft im breiteren Sinn und fasst darunter alle Arten von Cyber-Angriffen per E-Mail. Oft ist die Phishing-E-Mail dabei nur der erste Schritt, um Zugang zum System zu erlangen. Das Nachladen von Schadsoftware oder der Missbrauch von E-Mail-Postfächern zum Versand weiterer Angriffe folgt unter Umständen unbemerkt. Der technische Angriff kann dabei über geklonte Login-Seiten stattfinden oder Links nutzen, welche einen Drive-by-Download auslösen. Hier führt allein der Besuch einer Website zum Download einer Datei, welche im Anschluss gegebenenfalls Sicherheitslücken in veralteter Software ausnutzen kann. Cyber-Kriminelle können mit den neuesten technischen Standards durchaus mithalten: Da Dateitypen mit direktem Systemzugriff (wie beispielsweise .exe-Dateien) von E-Mail-Filtersystemen mittlerweile häufig aussortiert werden, nutzen mittlerweile 48% aller schadhaften E-Mail-Anhänge Microsoft Office-Dateien wie .docm oder .xlsm, welche über Makros Schadsoftware nachladen können (Symantec, 2019). Viele Browser sprechen außerdem eine Warnung aus, wenn sich der Nutzer auf nicht-verschlüsselten Webseiten („http") bewegt. Die Folge: 58% aller Phishing-Websites nutzen eine SSL-Verbindung, d.h. sind unter „https" erreichbar (APWG, 2019).

Da sich die Angriffsmuster dynamisch ändern, reichen generalistische technische Schutzmaßnahmen wie Firewalls oder E-Mail-Filter mit einfachen Heuristiken nicht mehr aus, um IT-Systeme effektiv abzuschotten. Cyberkriminellen steht eine Vielzahl an kostengünstigen Werkzeugen zur Verfügung, um mit geringem technischen Know-how komplexe Angriffe durchzuführen (Pienta et al., 2018).

Einen Schritt weiter als Phishing geht das sogenannte Spear Phishing, welches zielgerichtete Angriffe auf Personen oder Organisationen beschreibt. Die Kriminellen nutzen hier bestehende Vertrauensverhältnisse aus, indem sie sich auf Personen oder Sachverhalte beziehen, die der Empfänger bereits kennt. Beispiele sind E-Mails im Namen von Kollegen, gefälschte Rechnungen von tatsächlichen Lieferanten oder Anfragen, die mit Branchenwissen glänzen. Solche Angriffe nutzen oft Informationen aus öffentlich zugänglichen Quellen, im Fachjargon wird dies als Open Source Intelligence (OSINT) bezeichnet. Das Problem: Spear Phishing-

Angriffe sind heute nicht mehr mit großem manuellen Aufwand verbunden, sondern können automatisiert durchgeführt und millionenfach eingesetzt werden. Der altbekannte Glaube, Phishing-E-Mails erkenne man an Rechtschreibefehlern und fehlendem Kontext, ist für Nutzer im Arbeitsalltag demnach nicht mehr zutreffend.

5.2.2 Automatisiertes Spear Phishing in freier Wildbahn: Emotet

Das Jahr 2019 bewies eindrucksvoll, dass Spear Phishing kein Einzelfall mehr ist, vor dem sich nur hochrangige Ziele zu fürchten haben. Ein Beispiel für automatisiertes Spear Phishing im großen Stil ist Emotet. Die Schadsoftware versendet E-Mails mit schädlichem Dateianhang (häufig .docm oder .xlsm) oder Links und ist dabei in der Lage, „auf bestehende E-Mail-Konversationen zu antworten und daher authentisch wirkende E-Mails zu verschicken" (BSI, 2019). Dabei führt eine Erstinfektion dazu, dass organisationsintern weitere Phishing-E-Mails im Namen der Betroffenen versendet werden. Der eigentliche Schaden entsteht durch nachgeladene Software, beispielsweise durch Trojaner, welche den Tätern Komplettzugriff auf das Netzwerk verschaffen, bevor eine Ransomware eingesetzt wird. Diese verschlüsselt Daten oder ganze Netzwerke und fordert Lösegeld.

Die Schadsoftware hat Ende 2019 binnen weniger Tage für IT-Ausfälle bei Industrie und Bundesbehörden gesorgt (BSI, 2019), außerdem waren die Städte Frankfurt am Main und Bad Homburg, das Berliner Kammergericht, die Justus-Liebig-Universität Gießen und das Klinikum Fürth über mehrere Tage komplett offline (Heise, 2019a; Heise, 2019b; Heise, 2019c). Neben rein finanziellen Schäden brachten diese Angriffe Produktionsausfälle, die Abmeldung eines Klinikums von der Notfallversorgung und geschlossene Bürgerämter mit sich, sowie 38.000 E-Mail-Nutzer der JLU Gießen, welche sich neue Passwörter für ihren Account persönlich in der Turnhalle des Campus abholen durften. In den genannten Fällen fand Emotet durch das Aktivieren eines Makros in einem Dateianhang Zugang zum Netzwerk.

5.2.3 Verstärkte Gefahr durch Phishing im KI-Zeitalter

Neben Emotet sorgen auch andere automatisierte Spear Phishing-Angriffsmuster für immer schwieriger zu erkennende Phishing-E-Mails. Cyberkriminelle nutzen öffentliche Daten von Unternehmenswebseiten oder aus sozialen Netzwerken, um gezielte Angriffe zu generieren (Maedche et al., 2019). Diese OSINT-Analyse wird oft nicht mehr manuell durchgeführt – relevante Daten werden mithilfe von Crawling-Tools von Webseiten und aus Sozialen Netzwerken gesammelt und anschließend automatisiert zur Erstellung von maßgeschneiderten Phishing-E-Mails genutzt. Informationen wie die Namen der Geschäftsführung, firmeninterne

Strukturen oder Ansprechpartner sind für die Angreifer dabei genauso interessant wie persönliche Daten aus Sozialen Netzwerken, z.B. ehemalige Arbeitgeber, Kontakte, Hobbys oder der Geburtstag. Dies alles hilft, Angriffe so persönlich und glaubwürdig wie möglich zu gestalten. Neben Phishing nutzen Cyber-Kriminelle auch andere Angriffsvektoren, wie beispielsweise Telefon-Phishing. Mehrstufige Angriffe beinhalten das gezielte Sammeln von Informationen, das Aufbauen eines Kanals ins Unternehmen bis hin zum technischen Angriff. Ein Blick in Richtung Zukunft lässt ahnen, welches Ausmaß an Komplexität mit künstlicher Intelligenz (KI) gesteuerte Social Engineering-Angriffe erreichen können: Den technologischen Vorteil von Conversational Agents oder selbstlernenden Angriffsmustern werden sich auch Cyberkriminelle zu Nutze machen.

5.3 Security Awareness ist unabdingbar – nur wie erreicht man sie?

Aufgrund der steigenden Gefahr durch Social Engineering-Angriffe wie Spear Phishing beinhalten Informationssicherheitskonzepte im organisatorischen Umfeld immer öfter Maßnahmen zur Security Awareness. Der Begriff Security Awareness beschreibt das Ausmaß, in dem Mitarbeiter die Bedeutung von Informationssicherheit in ihrem Unternehmen sowie die Tragweite ihrer eigenen Sicherheitsverantwortlichkeit verstehen und dementsprechend handeln (ISF, 2007). Security Awareness ist dabei als dynamischer Prozess zu verstehen: Neue Angriffsmethoden und -vektoren stellen, wie am Beispiel Emotet erläutert, Mitarbeiter und Informationssicherheitsverantwortliche vor ständig neue Herausforderungen. Ein anpassungsfähiges Awarenesskonzept sollte daher ein dauerhafter und integraler Bestandteil jeder Unternehmenskultur sein (Kruger and Kearney, 2006).

Die Fachliteratur (z.B. Wright et al., 2014a) nennt in Bezug auf Security Awareness häufig die von Kahneman (2011) beschriebene Unterteilung des menschlichen Denkens in schnelles („System 1") und langsames („System 2") Denken: Während System 1 für erfahrungsbasierte oder automatisierte Informationsverarbeitung zuständig ist und dabei intuitive, schnelle Entscheidungen erlaubt, beschreibt System 2 das abwägende, rationale und analytische Verhalten, welches zur Bewertung größerer und langsamer Entscheidungen genutzt wird.

Der Netzaktivist und Hacker Linus Neumann (2019) erläutert, wie sich im Bereich Security Awareness fast alle eingesetzten Maßnahmen auf das Training von System 2 fokussieren: Schulungsmaßnahmen sind auf das rationale Denksystem ausgerichtet und sollen durch Wissensvermittlung und Checklisten dem Nutzer eine Hilfestellung zum Erkennen riskanter

Inhalte geben. Der Haken dabei ist, dass man sich im Kontext von Phishing-Angriffen nicht auf System 2 verlassen kann: Intuitive, schnelle und teils emotionale Handlungen (System 1-Denken) bestimmen häufig das Verhalten in der konkreten Situation. Somit ist es unabdingbar, System 1 gegen Phishing-Angriffe abzusichern, d.h. intuitive Handlungen als Teil der Sicherheitsmechanismen zu antizipieren. Von organisatorischer Seite kann dies durch technische und prozessuale Maßnahmen unterstützt werden (siehe Abschnitt 5.5). Der Nutzer selbst wird jedoch weiterhin in der Pflicht bleiben, sicherheitsbewusst zu handeln. Im Rahmen von Phishing kann dieses Verhalten neben konservativen Schulungsmaßnahmen wie beispielsweise E-Learning durch „Selbsterfahrung" trainiert werden. Die Erfahrung, als Nutzer selbst getäuscht oder „gehackt" zu werden, kann z.B. im Rahmen einer Phishing-Simulation stattfinden und dabei einen wichtigen „teachable moment" für einen Trainingseffekt in „System 1" bieten (Neumann 2019).

Verschiedene Forschungsbeiträge haben sich bisher in Form von Feldexperimenten dem Thema Phishing-Simulation gewidmet (beispielsweise Wright and Marett, 2010; Williams et al., 2018), und hierbei den Einfluss von Beeinflussungsmechanismen oder Verhaltensfaktoren auf die Empfänglichkeit des Nutzers gegenüber Phishing-Angriffen untersucht. Dieser Beitrag geht einen Schritt weiter und präsentiert ein Feldexperiment, welches im organisatorischen Umfeld den Einsatz von Security Awareness-Trainingsmaßnahmen, speziell bezogen auf Spear Phishing, untersucht. Hierbei wurden Schulungsmaßnahmen zur Wissensvermittlung mit einer Phishing-Simulation zur „Selbsterfahrung" kombiniert. Weiterhin führen wir eine Kennzahl ein, welche Security Awareness in Organisationen standardisiert messbar macht.

5.4 Feldexperiment: Messen und Trainieren der Security Awareness im organisatorischen Umfeld

Die im Folgenden vorgestellten Daten wurden in Kooperation mit der IT-Seal GmbH, einem Anbieter für Spear Phishing-Simulationen und Security Awareness-Trainings, erhoben.

Abschnitt 5.4.1 beschreibt das Projektziel und die Rahmenbedingungen. Die darauffolgende Präsentation von Methodik und Ergebnissen gliedert sich in zwei Teile. Teil I (Abschnitt 5.4.2) konzentriert sich auf die Messung der Security Awareness, in Teil II (Abschnitt 5.4.4) wird auf die eingesetzten Trainingsmaßnahmen und deren Wirkung eingegangen. Zur standardisierten Bewertung der Ergebnisse dient das Framework der Kennzahl „Employee Security Index", welches in Abschnitt 5.4.3 eingeführt wird.

5.4.1 Projektziel und Rahmenbedingungen

Das Feldexperiment wurde über einen Zeitraum von sechs Monaten in einer deutschen Organisation mit ca. 500 Mitarbeitern durchgeführt (Roethke et al., 2020). Ziel des Projekts war es, die Security Awareness in der Organisation zu messen und zu steigern. Dafür wurden von IT-Seal etablierte Schulungsformate wie E-Learning und Präsenzschulungen in Kombination mit einer realitätsnahen Phishing-Simulation unter Nutzung des „teachable moments" eingesetzt: In dem Moment, in dem der Nutzer durch Selbsterfahrung („ich werde gehackt") auf das Risiko von sorglosem Umgang mit E-Mails aufmerksam wird, entsteht eine erhöhte Lernbereitschaft.

Die Testgruppe besteht aus 511 Mitarbeitern einer deutschen Organisation. Von den 511 Personen sind 58% weiblich und 42% männlich. Alle Mitarbeiter nutzen E-Mail als tägliches Kommunikationsmittel. Die Durchführung einer „Phishing Awareness-Maßnahme" wurde innerhalb der Organisation ca. 3 Wochen vor Beginn des Projekts in Form eines Rundschreibens angekündigt. Um dem Mitarbeiter- und Datenschutz zu genügen, erfolgt die Auswertung der Messdaten auf Gruppenbasis mit einer Mindestgruppengröße von 30 Mitarbeitern.

5.4.2 Teil I: Messung der Verwundbarkeit gegenüber Spear Phishing-Angriffen

Methodisches Vorgehen: Spear Phishing-Simulation

Im Rahmen des Projekts standen je Mitarbeiter der Vor- und Nachname, die E-Mail-Adresse sowie Abteilung und Position in der Organisation zur Verfügung. Zusätzlich wurden von IT-Seal auf beruflich genutzten sozialen Netzwerken (Xing, LinkedIn) sowie auf der Webseite der Organisation öffentlich verfügbare Informationen über Organisation und Mitarbeiter gesammelt (OSINT-Analyse).

Die Menge dieser Informationen wurde genutzt, um zielgerichtete Phishing-Angriffe zu simulieren. Über den Projektzeitraum von sechs Monaten erhielt dabei jeder Mitarbeiter 2-3 E-Mails pro Monat, wobei Inhalt und Zeitpunkt individuell waren. Die technische Zustellbarkeit der E-Mails wurde im Rahmen des Experiments durch ein Whitelisting des Absenderservers gewährleistet, sodass im Penetrationstest der Faktor Mensch isoliert betrachtet werden konnte.

Die Auswahl der simulierten Phishing-Szenarien reichte dabei von generischen Angriffsversuchen (z.B. „Ihr Postfach ist voll") bis hin zu zielgerichteten Spear Phishing-Angriffen, welche einen oder mehrere der in Table 5-1 beschriebenen Parameter nutzten, um

die Glaubwürdigkeit des E-Mail-Szenarios zu steigern. Insgesamt standen für die Simulation ca. 80 Szenarien-Templates zur Verfügung, welche automatisiert individuell an den Empfänger angepasst wurden.

Parameter	**Beschreibung**	**Beispiel**
Anrede	Nutzung des Namens des Empfängers in der Anrede	Sehr geehrter Herr Mustermann / Hallo Max
Absender	Der Absender ist eine reale Person	Name einer Kollegin oder der Geschäftsführung
Domain	Die Domain eines enthaltenen Links oder der Absender-E-Mail-Adresse ist an die Empfängerin angepasst, z.B. ist die Domain der Organisation mittels Nutzung eines Buchstabendrehers oder einer Subdomain nachgeahmt („spoofing“)	cornelia.chefin@musterfrima.de https://intern.musterfirma.de-index.info/...
E-Mail-Signatur	Die E-Mail-Signatur des Absenders ist nachgeahmt	Die organisationsinterne E-Mail-Signatur wurde aus vorigem E-Mail-Verkehr oder von der Webseite übernommen
Geklontes Design	Die E-Mail enthält bekannte Logos oder Designs	E-Mail im nachgeahmten Design von Dropbox oder Amazon
Branchen-kontext	Der Inhalt bezieht sich auf branchentypische Inhalte	Rückfrage einer Krankenkasse an einen Mitarbeiter eines Krankenhauses
Zeitlicher Kontext	Der Inhalt passt im zeitlichen Kontext	Weihnachtliche E-Grußkarte im Dezember
Bezug auf Fachbereich	Der Inhalt bezieht sich auf den Fachbereich des Empfängers	Bewerbungsschreiben an einen Mitarbeiter aus HR
Bezug auf Information aus Sozialen Medien	Der Inhalt bezieht sich auf eine von der Empfängerin veröffentlichte Information	Anfrage mit Bezug auf ein Hobby, welches auf sozialen Netzwerken angegeben wurde, oder mit Bezug auf einen ehemaligen Arbeitgeber

Table 5-1. Parameter zur Steigerung der Glaubwürdigkeit der simulierten E-Mails

Jede simulierte E-Mail enthielt einen Link oder Dateianhang, deren Öffnen über das Nachladen eines Tokens gemessen wurde. Parallel zur Phishing-Simulation wurden Schulungsangebote zum Thema Security Awareness ausgerollt (siehe Abschnitt 5.4.4).

Ergebnisse: Beispiele simulierter Angriffe und deren Erfolgsraten

Im Folgenden sind beispielhaft drei der simulierten Phishing-Szenarien dargestellt. Figure 5-1 zeigt eine generische Phishing-E-Mail, welche in dieser Form millionenfach versendet werden kann. Um der E-Mail Legitimität zu verleihen, wird in der Absender-Domain die Domain der

Organisation nachgeahmt (hier beispielhaft verdeutlicht durch „@musterfrima.de“). Der Aufwand zur Vorbereitung einer solchen E-Mail beschränkt sich demnach im Wesentlichen auf das Registrieren einer entsprechenden Domain. Im Feldexperiment öffneten 26% der Empfänger (14/53) den enthaltenen Link.

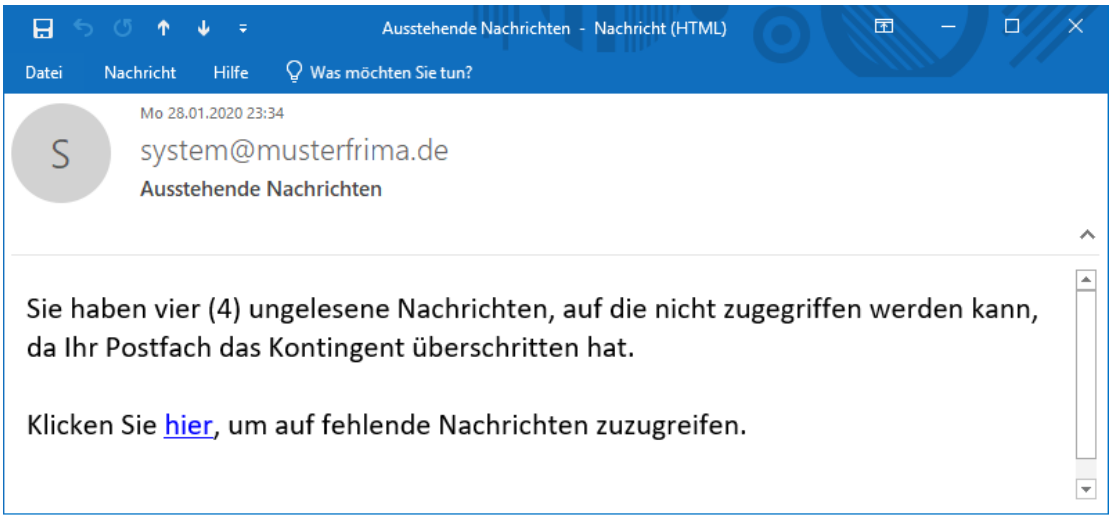


Figure 5-1. Simulierte Phishing-E-Mail „Ausstehende Nachrichten”

In Figure 5-2 ist ein Spear Phishing-Angriff dargestellt. Als Absendername wird der Name der Geschäftsführung genutzt, dieser lässt sich, wie auch die E-Mail-Signatur, ohne großen Aufwand auf der Webseite der angegriffenen Organisation finden und automatisiert verwenden. Der Link zeigt auf eine Domain, welche der Domain der angegriffenen Organisation täuschend echt nachgeahmt ist. Die E-Mail wurde 74 Mal versendet, der enthaltene Link wurde 33 Mal geöffnet (44%).

Im dritten Beispiel (siehe Figure 5-3) werden gezielt interne Strukturen ausgespäht und genutzt, um E-Mail-Verkehr zwischen Abteilungsleiter und Mitarbeiter zu fälschen. Der E-Mail hängt eine .docm-Datei an – hier ist bei realen Angriffen insbesondere das Öffnen das Makros mit einem sehr hohen Risiko verbunden. Im Experiment öffneten 23% der Empfänger (23/98) den Dateianhang, zwei Empfänger aktivierten im Anschluss das Makro.

Die drei dargestellten Beispiele gehörten im durchgeführten Experiment zu den „erfolgreichsten” Phishing-Szenarien.

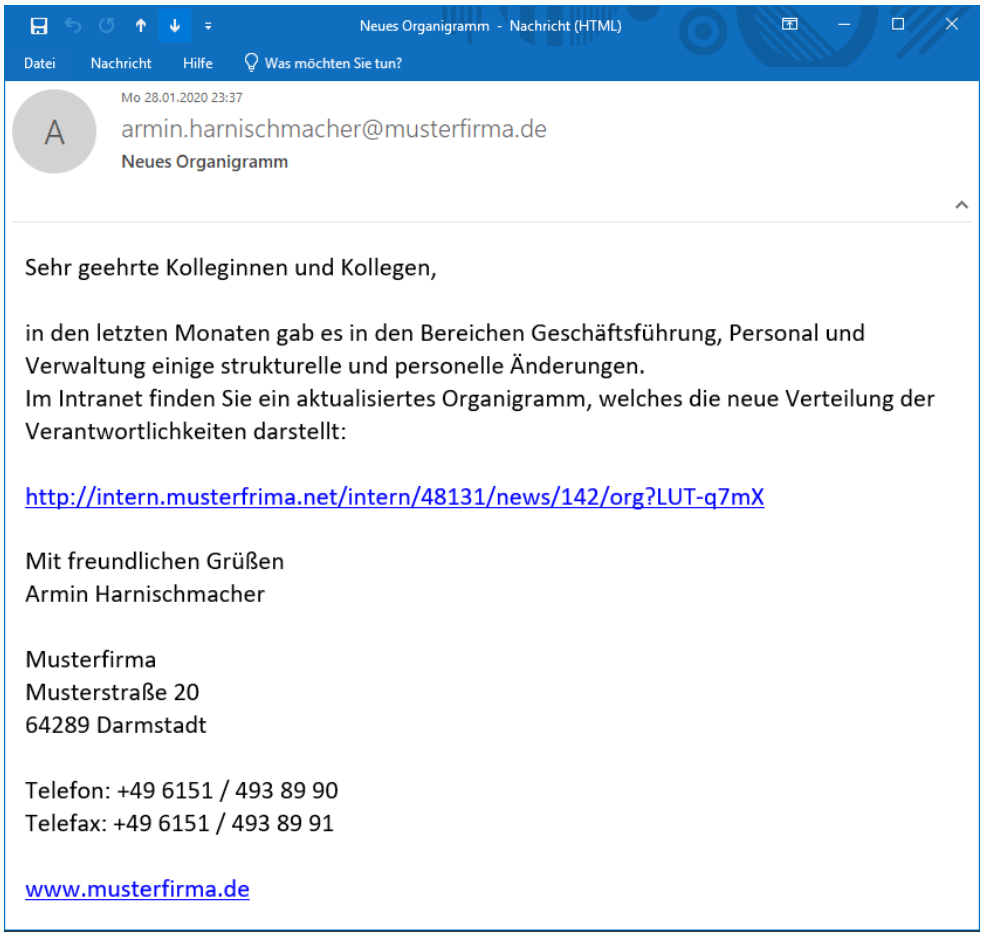


Figure 5-2. Simulierte Phishing-E-Mail „Neues Organigramm”

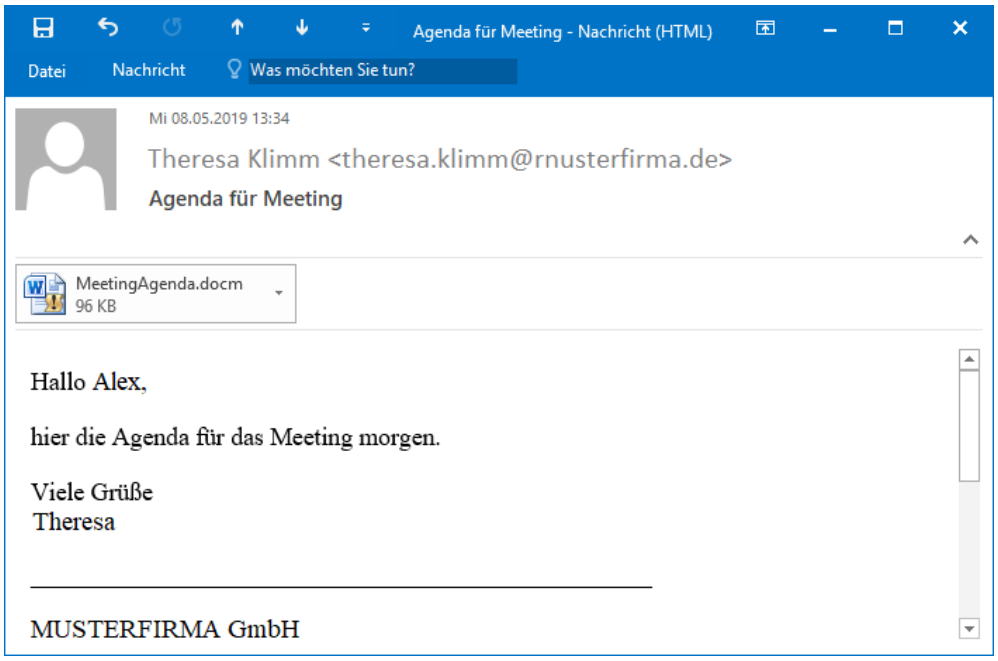


Figure 5-3. Simulierte Phishing-E-Mail „Agenda für Meeting"

5.4.3 Einführung der Kennzahl „Employee Security Index"

Um die Ergebnisse einer solchen Phishing-Simulation vergleichbar zu machen, betrachten wir die gemessenen Klickraten im Framework des „Employee Security Index" (ESI), welcher ein standardisiertes und reproduzierbares Verfahren zur Messung der Security Awareness darstellt (Franz, 2019). Wie oben beispielhaft dargestellt, können sich Phishing-Angriffe bezüglich ihrer Qualität und ihres Vorbereitungsaufwands stark unterscheiden. Wir teilen daher Phishing-Angriffe in drei verschiedene Level ein (siehe Table 5-2), welche sich am Vorbereitungsaufwand orientieren. Neben der technischen Vorbereitung, dem Klonen bestehender Designs, dem Erstellen von Malware oder dem Registrieren nachgeahmter Domains wird hier insbesondere die Zeit zur Informationsbeschaffung berücksichtigt. Der Angreifende wird hierbei als professionalisierter Cyberkrimineller eingeordnet, um die tatsächliche Gefahrenlage für Unternehmen und Organisationen möglichst realitätsnah abzubilden.

Level	Zeitaufwand	Beispiel
1	ca. 1h	Wenig vorbereitete E-Mail in Unternehmenssprache (Beispiel siehe Figure 5-1)
2	ca. 3h	Mäßig vorbereitete E-Mail, ggf. mit persönlicher Ansprache und Verwendung öffentlicher Informationen (Beispiel siehe Figure 5-2)

3	ca. 10h	Angreifer übernimmt in der E-Mail die Rolle eines Kollegen oder Vertrauten des Empfängers (Beispiel siehe Figure 5-3)

Table 5-2. Klassifizierung von Phishing-Angriffen

Die Kennzahl „Employee Security Index" nutzt die Daten einer Phishing-Simulation, um das Sicherheitsverhalten von Mitarbeitern bewertbar und vergleichbar zu machen. Auf einer Skala von 0 bis 100 definiert sich eine fiktive „vorbildliche" Nutzergruppe durch das Erreichen einer 90. Unter der Annahme, dass bei Phishing-Angriffen auf Unternehmen oder Organisationen eine Klickrate von 0% praktisch nicht erreichbar ist, definiert sich die „vorbildliche" Gruppe durch Toleranzwerte. Diese legen fest, welche Klickraten pro Level, d.h. pro für den Angriff aufgewendete Vorbereitungszeit, als ausreichend sicher bewertbar und dabei realistisch erreichbar sind. Der ESI berechnet sich für eine Testgruppe, z.B. Mitarbeiter einer Organisation, wie folgt:

$$ESI = \left(9 - \left(\frac{\sum_{k=1}^{3} A_L}{\sum_{k=1}^{3} n_L \cdot t_L} - 1\right)\right) \cdot 10$$

Hierbei ist n_L die Anzahl der simulierten Angriffe pro Level L, A_L die Anzahl der für dieses Level gemessenen Klicks und t_L der jeweils festgelegte Toleranzwert.

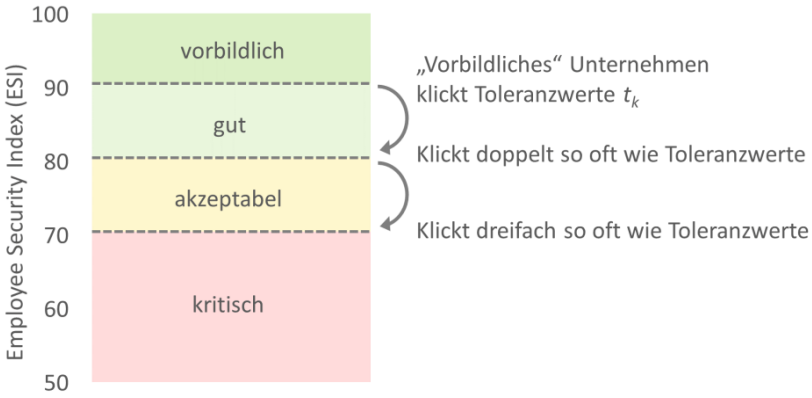


Figure 5-4. Bewertungsskala des Employee Security Index

Der ESI macht eine Testgruppe mit der als „vorbildlich" definierten fiktiven Gruppe vergleichbar (siehe Figure 5-4). Klickt die Testgruppe in einer vergleichbaren Simulation doppelt (dreifach) so oft wie die „vorbildliche" Gruppe, erreicht sie einen ESI von 80 (70). Die Skala ist unterteilt in die Bereiche „vorbildlich" (ESI ≥ 90), „gut" (ESI < 90), „akzeptabel" (ESI < 80) und „kritisch" (ESI < 70). Der ESI findet seit 2018 Anwendung in Security Awareness-

Projekten von IT-Seal. Basierend auf der Erfahrung mit kontinuierlichen Phishing-Trainings und dabei erreichtem Verhalten wurden die Toleranzwerte für einen ESI von 90 auf t_1=1,7%, t_2 = 4,1% und t_3 = 6,1% Klickrate festgelegt.

Das Thema „Security Awareness" ist selbstverständlich sehr viel breiter und kann nicht alleine durch das Thema Phishing Awareness beschrieben werden. Letztere eignet sich durch die konkrete Messmöglichkeit im Rahmen einer Phishing-Simulation jedoch stark zur Ermittlung eines Vergleichswerts, und wurde daher als Grundlage eines solchen Index herangezogen. Die Ausweitung der Messung auf weitere Security Awareness-Bereiche bietet eine umfassendere Bewertung und sollte Teil weiterer Forschungsmaßnahmen sein.

5.4.4 Teil II: Messung der Wirksamkeit des Security Awareness-Trainings

Aufbau des Security Awareness-Trainings

Die in Abschnitt 5.4.2 beschriebene Phishing-Simulation wurde als Teil eines Security Awareness-Trainings durchgeführt, welches aus drei Komponenten besteht.

a) Präsenzschulung

Vor Start der Phishing-Simulation absolvierte eine Auswahl von 100 der 511 Mitarbeiter eine ca. 90-minütige Präsenzschulung zum Thema „Phishing, Vishing, Human Hacking". Hier wurden die Themen Social Engineering, (Spear) Phishing, E-Mail-Sicherheit, Soziale Medien und Passwortsicherheit behandelt, aus den Medien bekannte Vorfälle besprochen und anhand eines Live-Hackings gezeigt, wie ein Phishing-Angriff ablaufen kann. Die Zuordnung, welche Mitarbeiter die Präsenzschulung absolvierten, stand für die weitere Datenauswertung im Verlauf des Projekts nicht zur Verfügung.

b) E-Learning

Als zweite Komponente wurde mit Start der Phishing-Simulation organisationsweit ein E-Learning ausgerollt, welches die unter a) genannten Inhalte in einem 30-minütigen Web-Based Training behandelt. Das E-Learning zeigt unter anderem auf, welches Risiko Spear Phishing birgt, und was bei in E-Mails enthaltenen Links und Dateianhängen beachtet werden sollte. Die Bearbeitung des E-Learnings war freiwillig, es wurden nach dem initialen Roll-out über sechs Monate vier Erinnerungs-E-Mails versendet. Insgesamt haben 377 der 511 teilnehmenden Mitarbeiter (74%) das E-Learning abgeschlossen. Sowohl E-Learning als auch die Präsenzschulung zielten auf das Training des System 2-Denkens ab.

c) Lernmoment im Rahmen der Phishing-Simulation

Als dritte Trainingskomponente diente die in Abschnitt 5.4.2 beschriebene Spear Phishing-Simulation selbst. Diese bietet einerseits die Möglichkeit zur „Selbsterfahrung" (das Gefühl, selbst „gehackt" zu werden) und damit eine Schulung des System 1-Denkens. Andererseits kann der Moment, in dem ein Fehler passiert (z.B. ein Klick auf einen gefälschten Link, oder das Öffnen einer risikobehafteten Datei) als „teachable moment" dienen, in dem eine besonders hohe Lernbereitschaft herrscht. Klickt ein Mitarbeiter auf einen in einer simulierten Phishing-E-Mail enthaltenen Link oder Dateianhang, so wird er zu einer interaktiven Lernseite weitergeleitet. Diese bietet am Beispiel der eben geöffneten E-Mail eine ca. einminütige Erklärung, wie die E-Mail als Phishing hätte enttarnt werden können. Besonderes Augenmerk liegt dabei auf dem Prüfen des Absenders der E-Mail sowie der Domain enthaltener Links, und der Vorsicht im Umgang mit Dateianhängen. Für alle Teilnehmer beginnt die Phishing-Simulation mit E-Mails des Schwierigkeitslevels 1 (siehe Table 5-2). Über den Projektzeitraum von sechs Monaten wird das Schwierigkeitslevel der E-Mails abhängig vom Klickverhalten des jeweiligen Mitarbeiters erhöht.

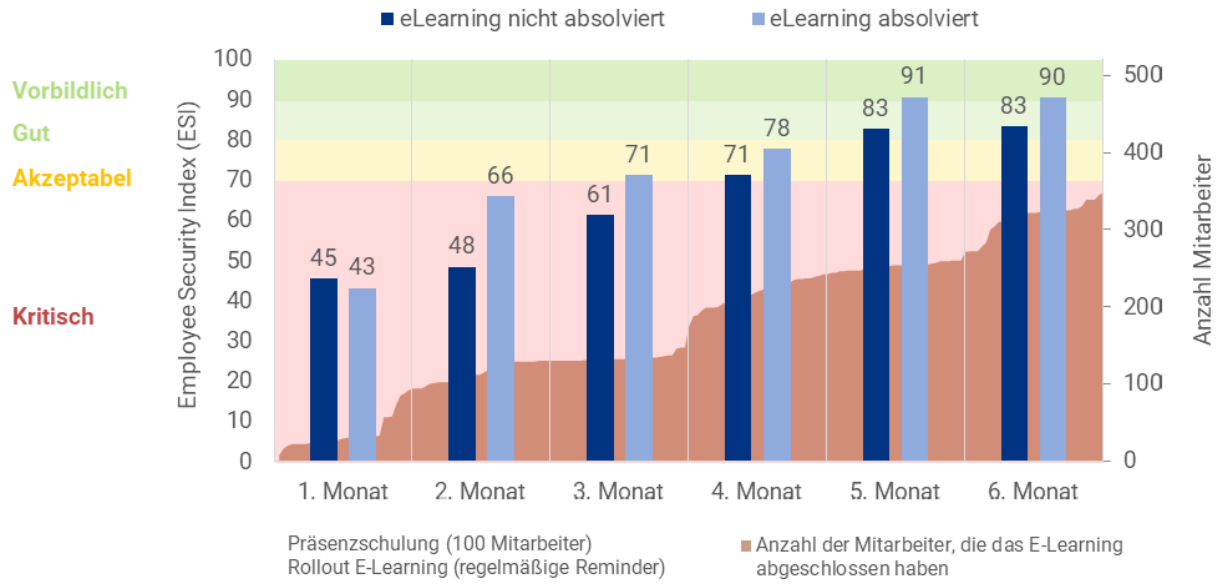


Figure 5-5. Trainingsverlauf des sechsmonatigen Security Awareness-Trainings

Entwicklung der Security Awareness im Trainingsverlauf

Figure 5-5 zeigt das Verhalten der Mitarbeiter gegenüber simulierten Spear Phishing-Angriffen in Form der Kennzahl „Employee Security Index" (siehe Abschnitt 5.4.3). Der ESI wurde organisationsweit von einem Startwert von 44 über sechs Monate hinweg auf 89 gesteigert. Dabei schnitten Mitarbeiter, die das E-Learning absolviert haben, im Durchschnitt besser ab als ihre ungeschulten Kollegen. Der Anteil der Mitarbeiter, die das E-Learning-Angebot nutzten und das 30-minütige Training komplett absolvierten, lag am Ende des Trainingszeitraums bei 67%. Auch bei den Mitarbeitern, die das E-Learning nicht absolviert haben, ist eine signifikante Steigerung des ESI messbar. Dies zeigt den Trainingseffekt der interaktiven Phishing-Simulation. In diesem konkreten Fall wurde von Seiten der Organisation entschieden, nach einer 6 bis 12-monatigen Pause die Maßnahme zu wiederholen – um den Effekt der Selbsterfahrung im Rahmen der Phishing-Simulation präsent zu halten, und um auch neue Mitarbeiter thematisch abzuholen. Da diejenigen Teilnehmer, welche auch das E-Learning absolvierten, deutlich bessere Ergebnisse erzielten, bietet es sich an, das E-Learning in einem weiteren Durchgang als verpflichtend zu gestalten.

5.5 Implikationen für Forschung und Praxis

Aktuelle Cyber-Angriffe auf Unternehmen und Organisationen zeigen, dass insbesondere Spear Phishing eine ernstzunehmende Gefahr ist. Aufgrund steigender Automatisierung wird die Menge solcher Angriffe in den kommenden Jahren stark zunehmen, und dabei für Nutzer schwieriger zu erkennen sein. Die Schadsoftware Emotet ist dabei ein prominentes Beispiel, welches Ausmaß an Komplexität und Schaden diese Angriffsmuster mit sich bringen. Aus dem in Abschnitt 5.4 beschriebenen Feldexperiment lassen sich folgende Implikationen ableiten, wie die Sicherheit gegenüber Cyber-Angriffen auf den „Faktor Mensch" gesteigert werden kann.

Multidimensionales Security Awareness-Training: Während die Vermittlung eines gewissen Grund-Know-hows (im vorgestellten Projekt umgesetzt durch ein E-Learning und Präsenzschulungen) für Nutzer nach wie vor als sinnvoll angesehen wird, reicht dies allein nicht aus, um eine nachhaltige Verhaltensänderung im Umgang mit Phishing-Angriffen zu bewirken. Das in dieser Arbeit beschriebene Feldexperiment zeigt, dass in Kombination mit wiederholter Selbsterfahrung in Form einer Phishing-Simulation unter Nutzung des „teachable moments" eine signifikante Verbesserung im Umgang mit Phishing-E-Mails erreicht werden kann. Um diese Art des Trainings durch Selbsterfahrung auch in anderen Bereichen der Security

Awareness zu ermöglichen, sollten künftige Forschungsbemühen genau hier ansetzen und beispielsweise die Bereiche Passwortkultur oder Vishing (Telefon-Phishing) in den Fokus nehmen. Für ein anhaltend hohes Sicherheitsbewusstsein sollten solche Trainingsmaßnahmen regelmäßig durchgeführt werden.

Nutzung von Kennzahlen: Zur Messung der Security Awareness in Organisationen kann eine Kennzahl wie der Employee Security Index (siehe Abschnitt 5.4.3) dienen. Idealerweise wird eine solche Kennzahl zum kontinuierlichen Monitoring der Security Awareness im strategischen (Informationssicherheits-) Management etabliert. Sie macht dabei Handlungsbedarf bestenfalls in Echtzeit erkennbar und bietet gleichzeitig die Möglichkeit der einfachen Kommunikation sowie Rückschlüsse auf den Return on Investment von Security Awareness-Maßnahmen. Von Management-Seite wird diese Kennzahl erfahrungsgemäß gut angenommen: IT-Seal setzt bereits mit mehreren Organisationen ein kontinuierliches Awareness-Programm um, wobei für einzelne Nutzergruppen abhängig von deren Rolle im Unternehmen ein „Ziel-Employee Security Index" festgelegt wird. Neben der Angriffssimulation werden dann weitere Schulungsmaßnahmen gezielt eingesetzt, um die Security Awareness auf den gewünschten Stand zu bringen und dort zu halten. Aus wissenschaftlicher Sicht ist die Erweiterung des hier vorgestellten ESI-Frameworks auf weitere Aspekte der Security Awareness von Interesse, um neben Phishing auch andere Bereiche standardisiert bewerten zu können.

Neben den Implikationen, welche sich direkt aus dem vorgestellten Feldexperiment ableiten lassen, sind aus Sicht der Autoren folgende Handlungsempfehlungen unabdingbar für eine Absicherung von Organisationen gegenüber aktuellen Cyber-Angriffen.

Ausbau technischer und organisatorischer Schutzmaßnahmen: Firewalls oder E-Mail-Filter sind mittlerweile gängige Maßnahmen im Kampf gegen Phishing, Malware und Co. Zusätzlich ist der Ausbau technologisch fortgeschrittener Schutzmaßnahmen, wie beispielsweise Advanced Threat Protection, stark zu empfehlen. Die bei Emotet-Angriffen häufig genutzten Makros stellen ein besonders hohes, und dabei schwierig zu bändigendes Sicherheitsrisiko dar. Makros können entweder organisationsweit deaktiviert oder nur mit digitaler Signatur erlaubt werden, um das Risiko einer Infektion mit Schadsoftware zu senken. Aus organisatorischer Sicht bilden etablierte Schutzmaßnahmen wie ein defensives Berechtigungsmanagement sowie regelmäßige Backups und Updates aller verwendeter Software eine unverzichtbare Sicherheitsgrundlage.

Umdenken nutzerverbundener Prozesse: Neben etablierten technischen und organisatorischen Schutzmaßnahmen sollten Prozesse so umgestaltet werden, dass intuitive Handlungen von Nutzern als Teil der Sicherheitsmechanismen antizipiert werden (siehe Abschnitt 5.3, „schnelles und langsames Denken"). Dies bedeutet, dass Sicherheitsmaßnahmen bezüglich des „Faktor Mensch" nicht lediglich in der Wissensvermittlung bzw. regelbasiert stattfinden sollten, sondern prozessual so ausgelegt sein müssen, dass auch intuitive Handlungen sicher stattfinden können. Ideen zur Umsetzung solcher Prozesse sind beispielsweise das Umdenken von der Kultur selbst wählbarer Passwörter hin zu unterstützenden Sicherheitsmechanismen wie Zwei-Faktor-Authentisierung oder neuen Standards wie dem passwortfreien Login FIDO2, sowie der Einsatz unterstützender Tools, welche Links im E-Mail-Programm oder Browser für den Nutzer transparent machen. Ein gut gemachter Phishing-Angriff kann für den Empfänger eine Stresssituation darstellen. Das Einschränken herunterladbarer Software auf vertrauenswürdige Quellen führt dazu, dass der Nutzer trotz intuitiver Aktionen nicht direkt risikobehaftete Dateiformate wie z.B. .exe-Dateien ausführen kann.

Sicherheitskultur etablieren: Zur wirklichen Umsetzung und Akzeptanz von Sicherheitsmaßnahmen gegenüber Spear Phishing in einer Organisation ist der Aufbau einer Sicherheitskultur unabdingbar. Dabei besteht die Herausforderung darin, ein potentiell negativ konnotiertes Thema (Informationssicherheit wird oft mit Angst, Frust oder Langeweile in Verbindung gebracht) für alle Mitarbeiter als relevant darzustellen und die Verantwortung des Einzelnen in der Unternehmenskultur zu verankern. Eine transparente und offene Kultur zum Umgang mit Fehlern ist dabei ebenso wichtig wie eine ausgeprägte Reporting-Kultur: Dringen neue Angriffe auf die Organisation schnell zur IT vor, kann hiervor gezielt gewarnt bzw. das Netzwerk technisch abgesichert werden. Gleichzeitig wird der Mitarbeiter aktiv in seiner Rolle als Mitverantwortlicher für Informationssicherheit eingebunden. Die dem Mitarbeiter zur Verfügung gestellte Meldekette sollte dabei möglichst schlank und aufwandsarm sein, am Beispiel von Phishing ist ein Melde-Button im E-Mail-Client denkbar. Weiterhin sind Ansätze in Richtung Gamification oder Belohnungssysteme vielversprechend, um diese sogenannten „extra-role behaviours" (Handlungen, die nicht in den eigentlichen Tätigkeitsbereich des Nutzers fallen), zu motivieren.

Zusammenfassend lässt sich festhalten, dass der Faktor Mensch in der Informationssicherheit auch in Zukunft eine entscheidende Rolle spielen wird. Informationssicherheitsverantwortliche stehen vor der Aufgabe, in der Belegschaft ein nachhaltiges Sicherheitsbewusstsein aufzubauen

und gleichzeitig nutzerverbundene Prozesse so umzudenken, dass schwerwiegende Fehler seltener möglich sind. Nur so können sich Organisationen auch gegen die künftig steigende Anzahl automatisierter Cyberangriffe wappnen.

Chapter 6: Information Security Culture in Times of Global Disruption

Title:	Facing Challenges Can Make You Stronger – How Global Disruptive Change Affects Organizations' Information Security Culture
Authors:	Anjuli Franz, Technische Universität Darmstadt, Germany Nihal Wahl, Technische Universität Darmstadt, Germany
Published in:	European Conference on Information Systems (2021), June 14-16, A Virtual AIS Conference

Abstract: An information security culture is the backbone of organizations' efforts to counter cyber attacks. The COVID-19 pandemic has fundamentally disrupted the way organizations and individuals work, with regard to remote working practices, new communication channels and top management's strategic decisions. Furthermore, new attack patterns exploit the vulnerabilities that come along with these changes. Based on interviews with 17 information security leaders, we formulate 10 propositions on novel facilitators and inhibitors for information security culture in times of disruptive change. Through the lens of punctuated equilibrium theory, we study which factors tip the scales for information security culture to radically transform in these unprecedented times. Our work contributes to both the research on organizational information security culture and the emerging body of literature on the impacts of the COVID-19 pandemic. In addition, we provide practitioners with valuable insights into crucial prerequisites for a strong information security culture.

Keywords: Information Security Culture, Disruptive Change, Punctuated Equilibrium Theory, Qualitative Study

6.1 Introduction

The COVID-19 pandemic has massively affected societies, organziations and individuals. In particular, it has substantially influenced the nature of work and the role that technology plays in the workplace (Carroll and Conboy, 2020). At the same time, cybercrime is on the rise, with criminals exploiting the lack of technological protection measures due to the sudden shift to remote work, and attacks continually evolving in response to changing situational factors (Naidoo, 2020). With 99% of attack attempts requiring human interaction, such as clicking on

a link or opening a file (Proofpoint, 2019), organizational information security culture hence becomes more important than ever.

Previous research has identified various factors that influence information security culture, and has presented instruments for its implementation and assessment (Da Veiga et al., 2020; Huang and Pearlson, 2019; Van Niekerk and Von Solms, 2010). Other researchers have argued that practice-oriented approaches predominate in information security culture research, and have called for focusing more on generating or testing theories to increase the maturity of this subfield of research (Karlsson et al., 2015). In the face of the current events, the impact of the COVID-19 pandemic on organizations has been investigated by recent works of information systems and management research, for example, with regard to organizational collaboration, communication and culture (Foss, 2020; Mithani, 2020; Waizenegger et al., 2020). Furthermore, IS research has portrayed its critical role in exploring how the behavioral and organizational aspects of newly emerging technologies can help to overcome this worldwide crisis (Ågerfalk et al., 2020). Information security culture is a highly complex construct that spreads across the organizational, leadership and individual level, and hence offers many leverage points for change. Furthermore, organizational information security meets novel challenges in the face of COVID-19: On the one hand, the pandemic causes an increase in cybercrime with ever-evolving social engineering attack vectors (Naidoo, 2020). On the other hand, the shift to remote work practices decreases information security leaders' oversight of employees' behavior. We argue that this makes information security culture an even more important asset of organizations. However, the impact of radical disruptive change (as currently presented by the COVID-19 pandemic) on organizational information security culture has received scant attention so far. With our work, we aim to fill this research gap by raising the following research question:

How does disruptive change affect organizations' information security culture?

Prior works have argued that organizational culture tends to be slow to change over time, and have highlighted the value of a longitudinal perspective (Cram et al., 2017). We have chosen a qualitative research approach, and have conducted 29 interviews with 17 information security leaders between June and October 2020. This study contributes to both IS research and the growing body of literature on the impact of the COVID-19 pandemic on organizations in multiple ways. Our findings reveal several external and internal factors that extend previous information security culture models. Furthermore, we find strong indications that novel short- and long-term influence factors emerge in the wake of disruptive change, yielding the

unexpected result that "facing challenges *can* make you stronger". In this context, we study our results through the lens of punctuated equilibrium theory in order to analyse why and how disruptive change influences information security culture. By formulating 10 propositions on information security culture in times of global disruptive change, we outline avenues for future research. Furthermore, our findings provide valuable insights for information security practitioners.

6.2 Theoretical Background

In this section, we first review pertinent literature on organizational information security culture. We present several frameworks that have been established by academia in close interaction with practice, and give a brief overview of topics that have previously been studied in this field of research. Afterwards, we address the research area of the COVID-19 pandemic, which has globally emerged in the beginning of 2020 and is rapidly disseminating to date (Bedford et al., 2020). We present recent research works on the impact of the pandemic on, for example, intra-organizational collaboration and communication. Lastly, we introduce the punctuated equilibrium theory as a theoretical lens for the analysis of the data gathered in our study.

6.2.1 Information Security Culture

Information security culture has been studied since the beginning of the twenty-first century (Vroom and von Solms, 2004; Schlienger and Teufel, 2003a) and constitutes a core construct of organizational information security research (Cram et al., 2017). Early literature has described information security culture as "including all socio-cultural measures that support technical security methods, so that information security becomes a natural aspect in the daily activity of every employee" (Schlienger and Teufel, 2003b, p. 1). Subsequent research has called for extending this end-user perspective, since information security is a management problem (Ruighaver et al., 2007). Since prior works have shown that top management support positively impacts employees' compliance with security policies (Hu et al., 2012), Ruighaver et al. (2007) argue that, whereas operational responsibility lies with middle management and end-users, it is crucial for top management (hence, the board or CEO) to visibly prioritize information security and incorporate security issues in organizational strategy (Ruighaver et al., 2007). In this work, we will use Da Veiga et al. (2020)'s extensive definition of information security culture, which includes, among others, regular communication and trainings, employees' attitude towards the protection of information assets, as well as management's

vision being aligned with the information security policy. It is important to note that, additionally to the expected in-role behaviour (such as following information security policies), employees' extra-role behaviour plays a vital role in information security culture. The latter consists, for example, in helping others with implementing information security policies, or offering comments intended to improve the organization's information security (Hsu et al., 2015).

Several research works have established comprehensive frameworks of organizational information security culture (Da Veiga et al., 2020; Van Niekerk and Von Solms, 2010; Van Niekerk and Von Solms, 2005; Huang and Pearlson, 2019; Da Veiga and Eloff, 2010; AlHogail, 2015). They build upon the model for corporate culture as presented by Schein (2010), which has become widely accepted amongst information security researchers. According to Schein (2010), culture is a property of a group, which begins to form whenever the group has enough common experience. To distinguish the structural elements of organizational culture, Schein (2010) has introduced a three-tier model. At the surface lies the level of *artifacts*, that is, visible and feelable structures and processes. In terms of password usage, for example, this could be "Do employees use different passwords for different accounts?". From outside the organization, the cultural level of artifacts is easy to be observed, but difficult to be deciphered. The next level are *espoused beliefs and values*, which describe strategies, goals and aspirations that exist in an organization. Getting back to password usage, these could be policies or procedures with regard to authentication. Thirdly, the underlying *shared tacit assumptions* build the deepest level of corporate culture. They are taken-for-granted, unconscious beliefs and values that are highly determinative of employees' behavior and perception, but difficult to be observed from the outside. In the password example, this could be "How serious are employees about using different passwords for different accounts?" (Schein, 2010; Van Niekerk and Von Solms, 2005).

Prior research has delineated organizational information security culture as a sub-culture of corporate culture. It has been argued that, for information security culture, *knowledge* has to be added as a fourth layer to Schein's model: while, for original definitions of corporate culture, it can be assumed that employees have the required know-how to perform their tasks, knowledge of information security is needed beyond that to perform those tasks in a secure manner (Van Niekerk and Von Solms, 2010). In the above example, this would correspond to employees knowing how to manage different passwords for different accounts (e.g., by using a password manager). Several other works have taken into account the practitioners' perspective, and have

established frameworks on how to implement an information security culture. Schlienger and Teufel (2003b) depict four stages, with top management commitment being the first stage, and organizational communication as well as training courses being followed by the commitment of the employees. In particular, programs for security education, training and awareness (SETA) have been identified to have a strong impact on information security culture (Cram et al., 2019; Haeussinger and Kranz, 2017). Other frameworks include the definition of culture change metrics, or rewards and punishments (Van Niekerk and Von Solms, 2005). Schein's model has been studied across the three levels of organizational behavior as presented by Szilagyi and Wallace (1983), and has been used to describe information security culture in the individual, organizational, as well as the leadership level. Furthermore, external influence factors, such as national culture or legal regulations, have been considered (Huang and Pearlson, 2019; Da Veiga et al., 2020). Turning to international standards, ISO 27001 provides requirements for training and regular updates in organizational policies for employees and contractors (ISO, 2013). Prior work has explored the role of various drivers for the diffusion of the ISO 27001 standard (Mirtsch et al., 2020).

Dhillon et al. (2016) have studied the effects of disruptive change in a prevalent security culture by means of a merger of two organizations, and have found the establishment of effective communication structures as well as a balance of informal, formal, and technical aspects to be essential in this context. Apart from that, the impact of organizational changes (e.g., digital transformation) on information security culture has received scant attention by prior literature so far.

6.2.2 The Impact of the COVID-19 Pandemic on Organizations

IS and management research have begun to study the impact of the coronavirus outbreak on organizations. Addressing the worldwide radical shift from on-site to virtual collaboration in March 2020, Waizenegger et al. (2020) have explored its impact on team communication and knowledge sharing. They have observed a change in organizational ad-hoc conversation, and have found that while daily stand-up meetings offer a means against isolation and loneliness, the overall amount of virtual meetings produces an overwhelming sense of intrusiveness for employees. Other works have discussed the rapid expansion of Enterprise Social Networks use and its influence on organizational rhythms (Dickinson, 2020), or the implementation of positive practices to avoid isolation (Gibson, 2020). Considering the potential impact on firms' organization design, Foss (2020) suggests that the COVID-19 pandemic is likely to leave permanent traces, for example in an increase in the modularization of tasks and task sequences,

which will lead to more formalization and a heavier use of individual-level rewards. Although there are several studies in the context of the COVID-19 pandemic with reference to organizations, research on the topic of information security culture is still missing in this context. Meanwhile, the attacker side has not been resting: Cyber criminals are exploiting the highly exceptional situation by making use of situational factors that lead employees to lower their guard, while selecting attack vectors that aim deliberately at users working outside the employer's normal security protections (Naidoo, 2020).

The above-described articles and further research show that the COVID-19 pandemic has had substantial implications for the nature of work (Carroll and Conboy, 2020). Given that technology is playing a key role in many aspects of the pandemic, such as allowing organizations to find new ways of working or to create new business models, IS research is presented with a broad research agenda: Calls have been made to explore to which extent current technologies can help overcome this worldwide crisis in the short term, and how organisations can utilize technology to recover in the long term (Ågerfalk et al., 2020; Adam et al., 2020). Looking beyond the technologies themselves, it is their behavioral and organisational aspects which will be challenging and critical for organizations to be resilient in times of crisis.

As for the theoretical lens, prior research has drawn, for example, on normalisation process theory to examine how emerging technology-driven practices can be embedded and routinized within an organization (Carroll and Conboy, 2020). Mithani (2020) has argued that existing theories of organizational adaption, which have been used for traditional environmental challenges such as economic or technological change, are inadequate in the face of life-threatening events such as natural disasters and pandemic diseases, since they do not account for the first-order consequences of physical and emotional threats which undermine operational continuity, routines, relationships, as well as the credibility of organizations' commitment to others.

6.2.3 The Punctuated Equilibrium Theory

The punctuated equilibrium theory has been used in IS and management research to characterize and investigate fundamental organizational change (Gregory et al., 2018; Guillemette and Pare, 2005; Romanelli and Tushman, 1994; Silva and Hirschheim, 2007). It is based on the idea that organizations are evolving through long periods of relative stability ("equilibrium"), where incremental change can be observed. This stability is interrupted by short revolutionary periods of radical change ("punctuation"), that disrupt established activity patterns and install the basis for new equilibrium periods (Romanelli and Tushman, 1994). IS

research has employed the theoretical lens of punctuated equilibrium theory to investigate, for example, the effect of IT consumerization on IT governance (Gregory et al., 2018), or the implementation of strategic information systems (Silva and Hirschheim, 2007). Romanelli and Tushman (1994) state that periods of revolutionary transformations are typically triggered by major changes in environmental conditions, substantial declines in the short-term performance, or significant change in senior management. There is no prior IS research that has used the punctuated equilibrium theory to study a phenomenon as radical and disruptive as the COVID-19 pandemic, which entails both a major change in environmental and corporate conditions as well as a substantial impact on organizations' short-term performance.

The three main concepts of punctuated equilibrium theory are organizational deep structure, equilibrium periods, and revolutionary periods (Gregory et al., 2018). The deep structure describes fundamental properties of an organization. Since strong interdependencies between its basic components make it resistant to transformation (Tushman and Romanelli, 1985), small, incremental changes in individual domains of organiziational activity will not accumulate to yield a fundamental transformation (Romanelli and Tushman, 1994). In periods of equilibrium, a system protects its deep structure and remains committed to underlying choices. Revolutionary periods are characterized by rapid and radical change, which dismantles an existing deep structure and establishes a new one (Gregory et al., 2018). In our work, we will employ the lens of punctuated equilibrium theory to interpret our inductive data on information security culture in the context of a global pandemic.

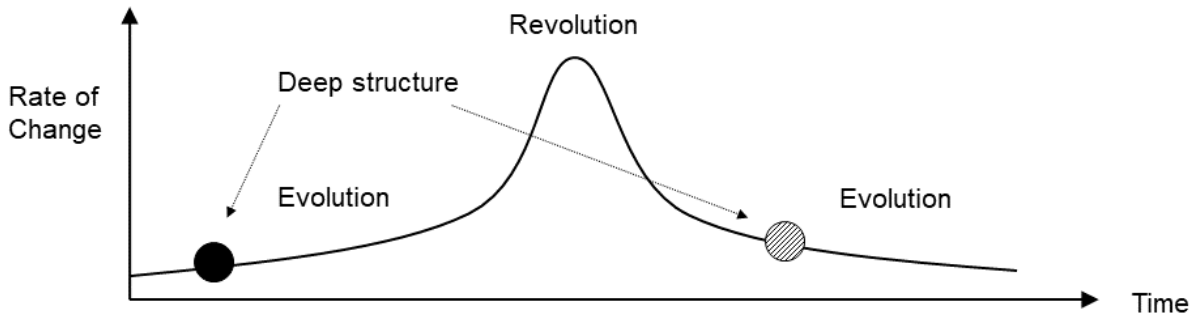


Figure 6-1. The model of punctuated equilibrium, based on (Silva and Hirschheim, 2007)

6.3 Methodology

6.3.1 Sample and Data Collection

The aim of this study was to explore the phenomenon of information security culture in the novel context of global disruptive change. Due to the exploratory nature of this topic, we chose

to conduct a qualitative study, which allowed us to make a detailed analysis of the relationships between the different factors and to consider contextual factors (Yin, 2017). Traditionally, qualitative studies have been proven to be a legitimate way to conduct research in IS literature, and have been employed by various prior research works on information security culture (Alshaikh, 2020; Dhillon et al., 2016) to describe phenomena and explore novel contexts. Accordingly, we decided to conduct semi-structured interviews. This allowed us to address the peculiarities of the respective interviewee's context, while ensuring that all the interviews covered the main topics. We used a standardized interview guideline, which was developed following recommendations by Yin (2017).

For our qualitative study, we conducted 29 interviews with 17 information security experts, who are employees of German or Swiss organizations. All interviewed experts are information security leaders in their organizations. In order to gain longitudinal insights into the interviewees' experiences, we aimed to interview each expert twice. The first interview round took place in June 2020. It focused on the experts' experiences and perceptions regarding information security culture within their organization before the COVID-19 pandemic, as well as on the implications of sudden change (e.g., with regard to remote work) during the first wave of the pandemic. The second interview round in October 2020 aimed at sharpening our understanding of long-term implications of the pandemic on information security culture. Since five experts were not available for a second interview, this yielded a total number of 29 interviews, with a total duration of 11 hours and 58 minutes. The interviews were held online in personal conversation. For easier analysis, they were audio recorded and transcribed.

As presented in Table 6-1, most of the interviewees have a proactive role and staffing-/budget-responsibility within their organizations. Several of them are chief information security officers (CISO), which constitutes the highest role in their business field. The respondents had practiced in this role for an average of nine years, ranging from 0.5 to 26 years. Most interviewees stated that, during the COVID-19 pandemic, a majority of their organization's employees were shifted to remote work.

6.3.2 Data Analysis Method

A time-related approach (pre-pandemic versus during pandemic) was employed to analyze the data. The aim of the data analysis was to retain and provide essential contents by abstracting a manageable collection of data that still illustrated a reflection of the data material. The methodological technique of "content analysis" is characteristic for this type of examination (Neuendorf, 2011; Weber, 1990). Since we did not have major theoretical assumptions in the

context of our research, we have used an inductive approach. Against this background, the categories of our results were not predefined or derived from existing theory, but were inductively derived from the transcribed interviews. Based on the content analysis technique and following the reducing code rules, the data material was reduced into an abstract form in order to paraphrase and generalize it by maintaining only the parts of substantial content, which was finally divided into categories (Corbin and Strauss, 2014; Mayring, 2004). For instance, the quotation of an interviewee *"Well, I'd say we had it relatively easy because we are such a critical infrastructure, which means we have to."* (I9) was coded – after a paraphrasing and generalizing process – to the category legal and regulatory requirements. As required, corresponding points in the material were assigned to the newly formed categories. To achieve reliability in our analysis, two independent persons coded and analyzed the data material by using the software tool MAXQDA 2020 (Richards, 2014). For each transcribed interview, codes were assigned to opinions that were found to be common amongst the participants by both persons separately. By merging these codes, we have combined all categories and marked only those that were coded by all.

ID	Position of the interviewee	Duration on this role in years	Industry	Employees in 2020
I-01	Responsible for information security awareness	1.5	Mechanical Engineering	11'500
I-02	Account Manager	4	IT Service	5'000
I-03	Responsible for information security awareness	2.5	Finance	50'000
I-04	Head of IT & Information Security Officer	26	Energy	280
I-05	Chief Information Officer (CIO)	19	Insurance	340
I-06	Information Security Officer	7	Social Insurance	100
I-07	Head of IT Infrastructure & Information Security Officer	10	Energy	2'000
I-08	Chief Information Security Officer (CISO)	1.5	Automotive Supplier	11'000
I-09	Information Security Officer	2	Food / Retail	26'000

I-10	Information Security Officer	20	Public Administration	1'200
I-11	Head of IT	17	Automotive Supplier	7'500
I-12	Employee in information security	11	IT Service	550
I-13	Information Security Officer	3	Finance	4'100
I-14	Responsible for cyber awareness	6	Public Administration	300'000
I-15	Chief Information Security Officer (CISO)	9	Transport and Logistics	22'000
I-16	Chief Information Security Officer (CISO)	13	Intergovernmental Organization	1'000
I-17	Security Awareness & Communications Officer	0.5	Telecommunication	18'000

Table 6-1. Overview of the interviewees and their organizations

6.4 Results

In order to illustrate our key findings in a comprehensive manner, we have ordered the presentation of the results around external (i.e., environmental) and internal (i.e., intra-organisational) influence factor dimensions that have been considered to affect information security culture by recent works (Da Veiga et al., 2020; Huang and Pearlson, 2019). On the internal side, we differentiate between the three levels of individual, organizational, and leadership factors. For each level, we have identified factors that act as either facilitators or inhibitors for information security culture. Regarding the timescale, we have further categorized the analyzed facilitators and inhibitors into three periods: (1) Factors that were predominant in the pre-pandemic era, (2) factors that prevailed during our first interview round in June 2020 (that is, right after the first wave of the COVID-19 pandemic in Germany), and (3) factors that dominated during our second interview round in October 2020 (that is, on the verge of the second wave). An illustrated summary of our results appears in Figure 6-2, where we present pre-pandemic factors on the left side, and pandemic-caused factors that supersede previous factors or newly add to prior concepts of information security culture on the right side.

Given the breadth of the topic, we illustrate factors that were found to be common among at least a number of interviewees. These factors are described in detail in the following sub-sections. Drawing on our results, we present 10 propositions that delineate how a disruptive period of radical organizational change, such as the COVID-19 pandemic, will influence organizations' information security culture.

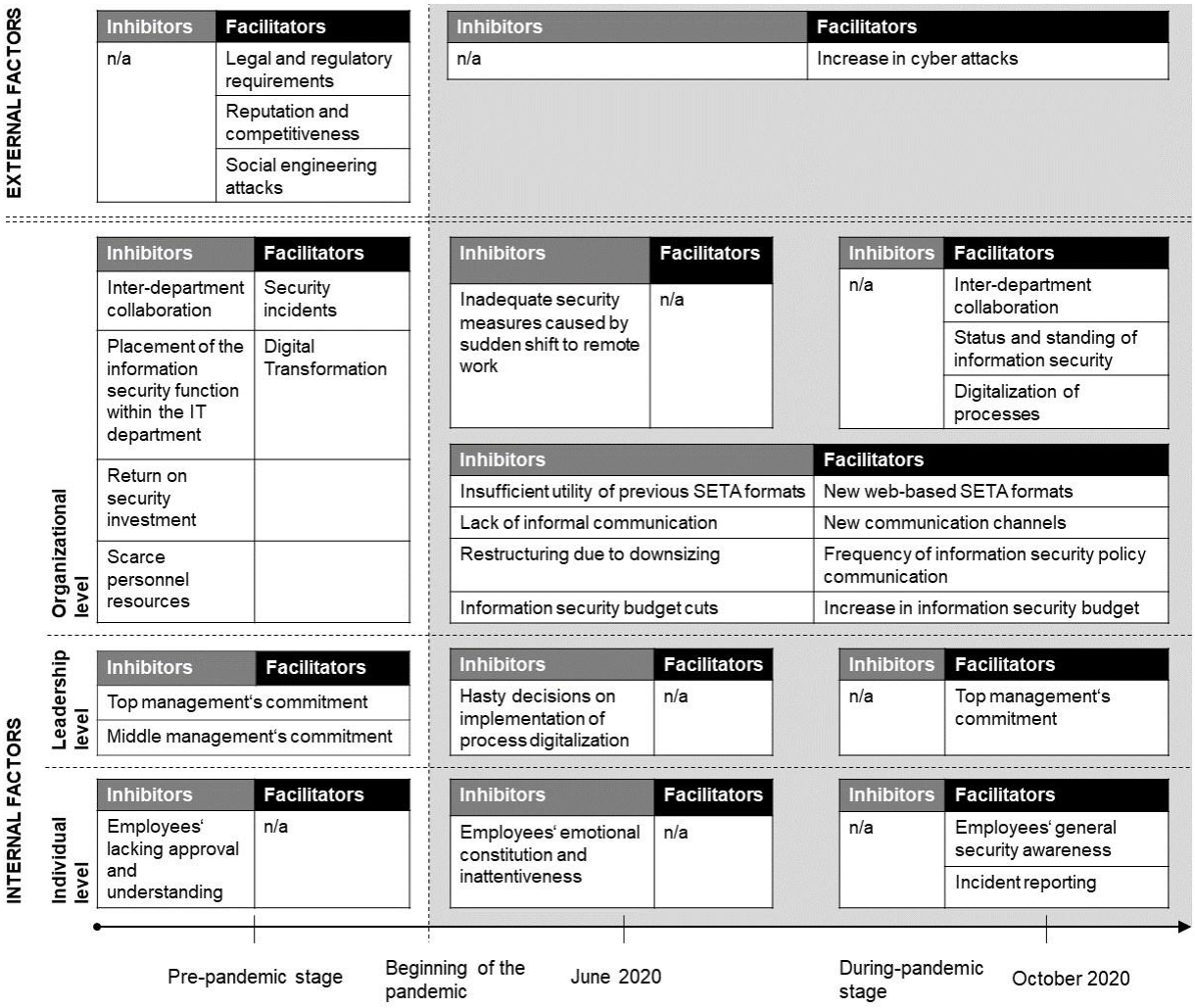


Figure 6-2. Inhibitors and facilitators of organizational information security culture before and during the COVID-19 pandemic

6.4.1 External Factors

With regard to external factors, we have found three relevant pandemic-independent factors that heavily influence organizational information security culture. First, we have identified **legal and regulatory requirements** as a key driver for the establishment of information security culture measures, with the majority of respondents stating that their organizations underlie regulations specifically concerning the "human factor" in information security. As one interviewee expressed it: *"Well, I'd say we had it relatively easy because we are such a critical infrastructure, which means we have to. We can't get out of it. I don't think otherwise the topic would be driven like that."* (I9). Furthermore, some respondents pointed out top management's personal liability as an important factor that accelerates decisions about information security, indicating that *"[…] as soon as somebody really has to take responsibility in writing, that's going to make it really fast."* (I7). Secondly, we found that **reputation and competitiveness**

constitute a substantial motivator for organizations to invest in their information security culture. One interviewed expert, an Information Security Officer of a large German bank, emphasized that with the following words: *"In the banking sector, it is a disaster if any account information can be read on a website. People would completely lose trust. A bank thrives on the fact that its customers entrust their money, as well as data and information to the bank."* (I3). In other industries, stacking up against competitors relies on information security, too, since customers increasingly demand information security standards as the following suggests: *"What always works, is if a customer demands, let's say, a standard or something like that. If that's one of our important customers, then money and resources will not be a problem."* (I8). Respondents also discussed the increased occurrence of **social engineering attacks** as a facilitator for the priorization of information security culture measures within their organization. As one interviewee expressed it: *"We realized that technical measures are often not effective if it's our employees who show improper conduct. Particularly with regard to ransomware, we were obliged to address the human factor."* (I10)

With the start of the pandemic, a majority of interviewed experts noticed a general **increase in cyber attacks** on their organization, stating they *"definitiely had more attacks"* (I10) as well as a *"sharp increase in phishing attempts"* (I1). They described that this factor, together with growing media attention on cyber attacks, added to the overall information security culture in their organization (e.g., I3, I11). In accordance with the above described factor, we formulate the following proposition:

Proposition 1: A global disruption (such as the COVID-19 pandemic) will lead to a noticeable increase in cyber attacks, which will in turn facilitate information security culture.

6.4.2 Internal Factors: Organizational Level

With respect to the pre-pandemic era, we have identified **poor inter-department collaboration** as a strong organizational inhibitor of information security culture. Information security officers reported other departments not respecting information security guidelines (I6), internal fights with data protection officers (I5), and, in particular, conflicts with the internal communications department: Interviewees described that they had to *"[…] fight hard to be allowed to use certain communication tools and formats […]"* (I15), and that they *"[…] did have a lot of security training materials, but were not able to disseminate them."* (I16). Furthermore, we have found that the **placement of the information security function** within the organization is a crucial factor for information security culture to thrive. In other words, organizations whose information security function was placed within the IT department reported the wish for an IT-

independent function, expecting to thereby ease collaboration with other departments and profit from non-technical staff's expertise (e.g., I4, I9, I16). Additionally, **scarce personnel resources** have been identified to restrain small teams from focusing on culture-related topics (e.g., I4, I15). In contrast, interviewees from organizations which did have an IT-independent information security function, stated this factor to be vitally important (e.g., I8, I13, I17). Many of them had direct reporting lines to top management (e.g., I1, I8, I12, I17). However, all respondents described measuring the **Return on Security Investment** as challenging, making it difficult to get top management's approval for information security budget. On the facilitator side, **security incidents within the organization** have been identified as an enabler for information security culture: *"They need to see with their own eyes that there are things that really hurt when they go wrong, and that's when the topic comes up to surface, even for management."* (I2). Furthermore, we have found **digital transformation** to be a major player in enabling information security culture. Interviewees have described digitalization of products and services as a catalyst for dealing with security topics (e.g., I1, I8), and stated that digital transformation *"[…] has started to drive a cultural change towards security by design […]"* with regard to projects and processes (I1).

Turning to the times of pandemic crisis, the overwhelming majority of respondents agreed on the fact that the very sudden introduction of remote working practices led to several problems in terms of information security. We have found that, during the first months of the pandemic, **inadequate security measures caused by the sudden shift to remote work** have significantly increased organizations' vulnerability towards cyber attacks. When interviewed again several months later, however, most interviewees stated that this issue had resolved itself because *"[…] things that had been allowed ad-hoc have now been reversed."* (I8). There were also comments about the **insufficient utility of previous SETA formats** due to the shift to remote work practices. Some interviewees whose organizations had not established web-based training formats before, stated that they were *"[…] at a loss with how to proceed […]"* (I16), with planned activities such as security games or security awareness days being cancelled due to the pandemic. Whereas organizations that were digitally more advanced also argued that *"[…] certain kinds of trainings just don't work that well online, unless you work with really small groups […]"* (I11), they also reported that the implementation of **new web-based SETA formats** has opened the door to new opportunities for fostering security culture, such as international lunchtime talks or higher coverage of training measures across the organization (e.g., I3, I10). We found that, when asked about intra-organizational communication, several

interviewees expressed concerns about a **lack of informal communication** due to decentralized work. One expert stated that *"since the 'grapevine' is severely restricted, we often just don't get the latest news on what's happening security-wise. Information security culture has been difficult before, this [the pandemic] makes it even harder."* (I9). On the other hand, the pandemic had given a push to implement **new communication channels**, for example, virtual collaboration tools or intranet platforms (e.g., I8, I10, I13). Without exception, all interviewees reported a rise in the **frequency of information security policy communication**, in particular with respect to remote work policies. As for the economical impact of the pandemic on information security measures, our interviewees' experiences were diametrically opposed. In the case of industries which were heavily impacted by the COVID-19-caused economical crisis, such as automotive or transport, some respondents said that their companies were **restructuring due to downsizing**, resulting in **information security budget cuts** as well as delays in budget release and decrease in information security personnel (e.g., I11, I15). Other respondents, coming from less affected industries, said that their department was *"[…] one of the few to not suffer from budget cuts or short-time work"* (I1), and some organizations whose economic situation was not affected by the pandemic even experienced an **increase in information security budget**, stating that *"information security even got granted higher budgets than usual, because of the pandemic"* (I10). We thus posit the following three propositions:

Proposition 2: A sudden shift to remote work will temporarily increase an organization's vulnerability towards cyber attacks.

Proposition 3: SETA programs and information security-related internal communication will undergo tremendous change in the context of a sudden shift to remote work and will therefore demand novel approaches.

Proposition 4: The impact of a global economic crisis on an organization's resources for information security culture will strongly depend on the organization's industry.

The analysis of the interviews that have been conducted more than half a year after the start of the pandemic has shown that things have started to settle within information security departments. We have found that three organizational factors have emerged from the pandemic context, which several interviewees perceived to act as powerful facilitators of information security culture: First, an **improvement in the inter-department collaboration** was mentioned by several respondents. One interviewee said that *"Thanks to the corona situation, the implementation of new security awareness campaigns went without resistance from top*

management, data protection officers, HR or the legal department – which was definitely unexpected." (I1) The CISO of an international automotive supplier echoed this view with the following words: *"We have established a task force with HR, logistics, IT, security and others, which first aimed to ensure business continuity during the pandemic. Now we use it to optimize 'cyber resilience' […], where we don't differentiate between the pandemic or hacker attacks, but look at the whole system from a higher level. […] This has extremely helped information security issues to pick up speed."* (I8). Secondly, several interviewees felt a significant change in the **status and standing of information security** within their organization. One interviewed person indicated that *"the topic of information security remarkably gained momentum across the company"* (I3). Another respondent illustrated this point clearly by describing how *"problems that have always been perceived as side issues by everyone outside information security, now suddenly have a different status."* (I1). Thirdly, the pandemic has been described as a significant contributor for the **digitalization of processes** due to the shift to remote working practices (e.g., I6, I13), and thereby as an implicit facilitator of information security culture: *"Corona has directly given a boost to digitalization, and therewith indirectly to security, that's how I'd say it."* (I13). As shown above, new emerging factors that are grounded in the pandemic context can strongly influence information security culture in the long term. Hence, we posit that:

Proposition 5: A disruptive period of radical change can yield a sustainable enhancement of inter-department collaboration, which will facilitate organizations' efforts to build an information security culture in the long term.

Proposition 6: Periods where information security is at the same time endangered and crucial across the organization as a whole will lead to an advancement of information security's status and standing within the organization.

Proposition 7: A sudden shift to remote working practices will yield a progress in digitalization, which will in turn implicate progress in information security.

6.4.3 Internal Factors: Leadership Level

Top management's as well as **middle management's commitment** to information security has been found to play a key role in information security culture across all interviewed organizations. Interviewees have described management's priorization of information security issues and participation in training and awareness measures as a multiplier for organizational information security at large (e.g., I4, I10, I15). With regard to the pre-pandemic era,

interviewees' experiences with top management's commitment varied widely. Approximately half of the respondents stated that they were very satisfied with top management backing information security-related topics by all means, describing a *"[…] significant increase in management's risk awareness over the previous years"* (I14) or acknowledging that they had *"[…] reached an absolutely amazing level of top management support"* (I7). Many interviewees mentioned that they had implemented specific training modules for management (e.g., I16) or held security awareness presentations for all members of the board (I15). With respect to middle management, the interviewees agreed in stating that leaders' commitment depends highly on their individual openness to the topic of information security (e.g., I5, I15).

With regard to the pandemic, we found that some interviewees complained about top management's **hasty decisions on the implementation of process digitalization**, as illustrated by one expert: *"[…] and then you're left to mop up the mess […]"* (I9). Towards the second wave of the COVID-19 pandemic, however, we observed a significant change in several interviewees' perception of **top management's commitment**. Respondents described that *"[…] decisions that had always been postponed have now been made […]"* (I10), that top management had made information security training mandatory for all employees or even published statement videos, personally taking a stand on security (I1). Those interviewees whose organizations were sorely afflicted by the economic crisis stated that, even though budget and personnel were scarce, they felt supported in security issues that they rated as critical (e.g., I11, I15). However, we did not observe a change in middle management's commitment. According to these findings, we formulate the following proposition:

Proposition 8: In times of global crisis (such as the COVID-19 pandemic), top management's risk awareness will lead to an increase in their commitment to information security.

6.4.4 Internal Factors: Individual Level

As for the individual level, we asked the interviewed experts about their perception of individual factors influencing information security culture in their respecitve organization. For the pre-pandemic era, we found that a majority of interviewees rated **employees' lacking approval and understanding** of information security-related topics as a strong inhibitor (e.g., I5, I13). They stated that their main goal culture-wise was to generate awareness and acceptance among employees, for example by trying to transfer information security in the employees' personal context (e.g., I3, I13, I17). With the beginning of the COVID-19 pandemic, we found that many of the respondents were concerned with **employees' emotional constitution and inattentiveness.** One interviewee commented: *"I think that due to the exceptional, emotionally*

trying and frustrating situation of the pandemic, people are absent-minded and inattentive. I believe that they are more prone to potential cyber attacks." (I1) Others stated that they believed employees to act more carelessly (I9), since *"[…] when working from home, nobody is standing behind their back and watching their screen."* (I8).

When interviewed again later, we found that most experts' perceptions had changed: They felt that **employees' general security awareness** had increased. They ascribed this to prior information security training, to information security policies being frequently communicated (e.g., I3, I8), and to the fact that employees understood their own responsibility (e.g., I10, I11). One interviewee stated: *"With the whole situation deepening due to corona, and our firm facing downsizing and the like, people are becoming more conscious that things could get even worse if they do not support information security."* (I1). Additionally, several respondents reported that employees' **reporting of potential incidents**, such as phishing e-mails, had increased (e.g., I10). This indicates that individual factors of information security culture strongly depend on situational factors. We hence posit:

Proposition 9: The start of a disruptive crisis will be emotionally demanding to employees in general, and will thereby decrease their attentiveness towards cyber attacks.

Proposition 10: In lasting times of crisis (such as the COVID-19 pandemic), employees will understand their role in their organization's resilience against cyberattacks, and will therefore become more information security-conscious.

6.5 Discussion

The COVID-19 pandemic is arguably the most defining global crisis that we have witnessed in the past decades. In particular, the pandemic has proven to be a crucible for workplace transformation (Dickinson, 2020; Waizenegger et al., 2020). Drawing on our qualitative study as presented in the previous sections, we have deduced 10 propositions that illustrate how a disruptive period of radical change, such as a global pandemic, will influence organizations' information security culture.

In summary, our findings show that in the wake of the pandemic (1) the radical change of where and how employees work, (2) the rise of cybercrime, as well as (3) the economic uncertainty have led to substantial changes in all three cultural dimensions as presented by Schein's (2010) model. Whereas novel artifacts have emerged, for instance, in the form of improved inter-department collaboration or top management's personal statements on information security, espoused beliefs and values have been transformed, for example, with regard to renewed

information security policies and the frequent communication thereof. Furthermore, a change in the underlying level of shared tacit assumptions has been identified in terms of employees' general security awareness as well as the status and standing of information security within the organization at large.

The results of our study confirm important factors that have been found to influence information security culture by prior research works, such as **SETA programs** (Da Veiga et al., 2020; Schlienger and Teufel, 2003b) or **inter-department collaboration** (Huang and Pearlson, 2019). Furthermore, our study reveals several short- and long-term factors that influence information security culture, that have not been considered by previous resesarch models. Regardless of the COVID-19 pandemic, we have identified the **placement of the information security function** within the organization to play a crucial role, with a placement within the IT department hampering the establishment of an information security culture, for example by inhibiting inter-department collaboration. Second, we have identified **digital transformation** as a long-term facilitator for organizational information security culture at the organizational level. The digitalization of products and services acts as a catalyst for dealing with security topics, and drives a cultural change towards security by design. During the pandemic, progress in **digitalization of processes** due to a sudden shift to remote work has implicitly augmented information security culture. Third, we have found that the behavior of the attacker side is an important external factor affecting information security culture. With regard to the pandemic, a sudden **increase in cyber attacks** has been found to advance information security culture at all three (organizational, leadership and individual) internal levels. Fourth, we have observed that the **status and standing of the information security function** within an organization has a significant impact on information security culture, and that it can change promptly in the face of disruptive change. Lastly, while the employees' emotional condition has been taken into account by prior works in terms of working time or working atmosphere (Da Veiga et al., 2020), the sudden implications of a life-threatening global event such as the COVID-19 pandemic for **employees' emotional constitution and attentiveness** had not been considered so far. In the remaining of this section, we will discuss our findings through the lens of punctuated equilibrium theory. We will then provide an overview of our contributions for theory and practice, and delineate avenues for future research.

6.5.1 Information Security Culture Through the Lens of Punctuated Equilibrium Theory

Our study has given insight into how information security culture changes amid global disruptive change. Culture is one of the most deeply entrenched, often unconscious parts of an organization (Schein, 2010). According to the approach of punctuated equilibrium theory, the "deep structure" (see Figure 6-1) of an organization consists of strong interdependencies between its basic components, which make it resistant to transformation (Tushman and Romanelli, 1985). In the case of information security culture, this deep structure lies, for example, in the priorizitation of information security within the organization's strategic goals (e.g., in the form of **top management's commitment**, the **placement of information security within the organization**, or available resources), or in the shared, tacit assumptions about the **standing and status of information security** (e.g., by means of **inter-department collaboration** or **employees' general security awareness**). In our study, we have found that the deep structure's basic components are not only strongly interdependent with respect to the pre-pandemic status quo, but can also heavily influence each other when disruption occurs. However, whereas some information security leaders have reported that their organization's information security culture has underwent a "period of revolutionary change" since the beginning of the COVID-19 pandemic, either towards "good" or "bad", others have not. In the following, we employ the punctuated equilibrium theory to discuss which factors tip the scales for information security culture in times of disruptive change, and exemplarily describe organizations at the respective ends of the spectrum. We thereby consider the beginning of the pandemic to be a "revolution" (stage 2), and the pre-pandemic and during-pandemic stages (stages 1 and 3, respectively) an "evolution" (see Figure 6-1).

The information security leaders I1, I8 and I10, for example, have described their organizations as relatively mature with regard to digital transformation. For instance, they have steering committees for topics regarding the digitalization of products and processes. I1 and I8 stated that their information security function is placed outside of the IT department. All three interviewees were satisfied with their **top management's commitment**. Furthermore, digital **SETA programs** had been implemented long before (I8, I10) or at the verge (I1) of the pandemic. I1 and I8 have described **inter-department collaboration** as a major inhibitor for information security culture in the pre-pandemic era. Although COVID-19 has relatively strongly affected both organizations economically, both I1 and I8 have stated in October 2020 that the pandemic has facilitated their efforts to build an information security culture. We have

found that **inter-department collaboration** and **employees' general security awareness**, both crucial influence factors of information security culture, have shifted from the inhibitor to the facilitator side (I1, I8). Furthermore, the information security department has been spared from **budget cuts and downsizing** (I1). Interviewee I10, whose organization was not affected economically, stated that **information security budget had been increased** in consequence of the pandemic. When interviewed in October 2020, all three experts (I1, I8 and I10) perceived the joint influence of the **increase in cyber attacks**, the visibility of **top management's commitment** as well as prior **SETA programs** to have led to an increase in both **employees' general security awareness** and **incident reporting**, as well as an overall enhanced **status and standing of information security** within the company. We therefore argue that, given a certain digital maturity, prior **SETA programs** as well as **top management's commitment**, a disruptive challenge such as the COVID-19 pandemic can indeed "make you stronger" and dismantle the existing "equilbrium" towards a more resilient information security culture.

In contrast, interviewee I15, an information security leader in the heavily affected air transport industry, stated that the economic effects of the pandemic have been so fundamental that **budget cuts** and **downsizing** were strongly inhibiting information security culture. On a different note, interviewee I16 stated that, while budget was not a problem, the organization's **digital immaturity** (in particular the lack of infrastructure for remote work) heavily inhibited information security culture in the face of the pandemic. We therefore argue that the above-described effect will not apply to organizations that lack digital maturity and infrastructure, or are located in industries that are heavily affected by disruptive economic change. On the contrary, a radical challenge will disrupt the existing "equilibrium" towards a weaker information security culture.

We lastly look at I17, whose organization we identified to be a pioneer in information security culture, with cultural values such as security by design being lived and regular SETA programs being a mandatory part of employees' work. We found that I17's organization has not experienced disruption by means of the COVID-19 pandemic. This leads to the assumption that information security culture can reach a saturation point, which makes it resilient and stable in the face of radical change.

6.5.2 Contributions to Theory and Practice

Our work contributes to research in several ways: (1) We extend prior models of organizational information security culture by several factors, such as the **placement of the information security function within the organization** or the **overall status and standing of information**

security, and hence contribute to an enhanced understanding of how organizational, leadership and individual factors influence information security culture. (2) Our work adds to the previous literature on the impact of radical workplace transformation on organizations, by studying information security culture in the context of the global COVID-19 pandemic. Drawing on interviews with information security leaders in June and October 2020, we find strong indications that novel short- and long-term influence factors emerge in the wake of disruptive change: At the beginning of the pandemic, **inadequate security measures** and **hasty decisions on digitalization** caused by the sudden shift to remote work together with **employees' inattentiveness due to their emotional constitution** yielded a temporarily weakened information security culture. In the long term, however, we find several pandemic-caused factors that fundamentally facilitate information security culture in the long term (see Figure 2 in Section 4). Based on our qualitative insights, we formulate 10 propositions that can serve as starting points for future research. (3) We employ punctuated equilibrium theory to analyze a real-world disruption that is more fundamental than any other context that has been investigated by IS research before. We find that the theoretical lens of punctuated equilibrium offers an eye-opening view on why and how disruptive change can influence information security culture: If certain prerequisites, such as the organization's **digital maturity** and **economic stability**, are met, organizations that were resistant to fundamental transformation before can experience a radical change in the context of a global disruption. Furthermore, our analysis shows that information security culture can reach a saturation point, which makes it resilient to periods of radical change.

Our study furthermore reveals valuable insights for practice. We highlight several factors that emerged as crucially important for an information security culture, and illustrate how tweaking, for example, **inter-department collaboration** or **top management's commitment** can yield great improvement. These insights can help information security practitioners to build a strong information security culture, both in times of disruption and equilibrium.

6.5.3 Limitations and Directions for Future Research

Like all research, our contributions are limited by the choices made in the design of our study. We interviewed 17 information security leaders from different industries and organization sizes in order to tap a variety of experiences. Our reliance upon individual reports however limits our ability to confirm the objective outcomes of our study. We therefore suggest further research to approach this topic with quantitative methods, where the formulated 10 propositions can be taken as a starting point for further investigations. Furthermore, we solely rely on the

perspective of information security experts, whereas information security culture spans across all levels of organizational behavior. We hope that further research will acknowledge, for example, the employee or top management perspective. The period of time in which our interviews took place constitutes another limitation of this work. While we have studied the development of information security culture until the verge of the second wave of the COVID-19 pandemic in Europe, further research is needed to investigate post-pandemic implications.

Chapter 7: Contributions and Conclusion

Both in the individual and organizational context, issues related to security and privacy are deeply ingrained within users' everyday social and emotional experiences. However, our understanding of these factors is still limited. As such, this thesis was motivated by the aspiration to unravel how and why socio-emotional factors influence individual and organizational information security and privacy, while simultaneously providing a deeper understanding of how these insights can be utilized for designing effective security and privacy-enhancing tools and interventions. Against this backdrop, five studies have been conducted and published in renowned IS outlets, using a variety of methodologies including literature review, qualitative interview study, online vignette experiment, and field experiment. Each study contributes to answering the overarching research questions of this thesis and examining the socio-emotional drivers of information security and privacy.

7.1 Contributions to Research

This thesis was guided by four overarching research questions addressing the human dimension of information security and privacy in general, and socio-emotional factors in particular. In the following, the contributions of this thesis are structured along these research questions.

RQ1: How do socio-emotional motives affect individuals' decision-making with regard to information security and privacy?

Articles B and C contribute to research on individual decision-making with regard to information security and privacy. Whereas a large body of literature has investigated users' reliance on systematic and heuristic processing to assess situations related to security and privacy (e.g., Luo et al., 2013; Dinev et al., 2015), it has largely limited its scope to a cognitive lens. With few exceptions (e.g., Bélanger and James, 2020; Burns et al., 2019a), IS research therefore lacks an understanding of how users' information and privacy decisions are influenced by socio-emotional factors. This thesis addresses this shortcoming and contributes to IS literature by demonstrating that social factors, such as users' perception of others' salience within their own socio-technical environment, and emotional factors, such as hedonic motivation to experience a feeling of "warm glow", are pivotal mediators in stimulating users' security and privacy behavior. In particular, socio-emotional factors do not only influence users' information disclosure behavior, but also their proactive protection of information assets. This is an important contribution to information security literature seeking to harness the

potential of human capacities in complementing technological security measures (Heartfield and Loukas, 2018; Zimmermann and Renaud, 2019), since it uncovers new ways how these capacities can be activated. In addition, this thesis responds to a call for new measurement tools to capture information security-related emotions (Renaud et al., 2021). By drawing on related fields such as donation literature (article C) and developing a scale reflecting the conceptual 3R framework (Kamleitner and Mitchell, 2019) (article B), this thesis provides measurement instruments that allow to zoom into the micro-level of users' socio-emotional processing and can serve as a basis for future studies in this field. Taken together, this thesis sheds light on socio-emotional factors as crucial drivers of information and privacy behavior that policymakers and tool designers can utilize to assist users with their security and privacy decisions. This leads to the second research question:

RQ2: How can information security and privacy tools and interventions leverage individuals' underlying socio-emotional motives to improve their security and privacy behavior?

Acknowledging the vital role of the user in preserving security and privacy, prior works have started to investigate tools and interventions that assist users in fulfilling this task (e.g., Volkamer et al., 2017; Jensen et al., 2017b; Schuetz et al., 2020). Drawing on the context of phishing as a proxy for prevalent security decisions, article A is the first to provide a taxonomy of user-oriented anti-phishing interventions. By systematizing prior knowledge and critically examining the different intervention approaches, this thesis is able to carve out crucial shortcomings and provide valuable suggestions for future works. Additionally, this thesis utilizes knowledge on socio-emotional drivers of user behavior from *RQ1* to investigate two novel stimuli for enhancing information security and privacy. Specifically, it reveals that an intervention that emphasizes the presence of others' personal data within the user's socio-technical environment significantly reduces instances of interdependent privacy infringements among users (article B). This highlights the importance of providing users with transparent and comprehensible information, as previously argued by prior works (e.g., Reuter et al., 2022). Furthermore, this thesis demonstrates that providing a cyber incident reporting tool that evokes a sense of personal satisfaction, known as the "warm glow" effect, significantly increases the usage of such reporting tools (article C). These findings are important because they unlock a new perspective on how tools and interventions can effectively stimulate secure and privacy-aware behavior. Taken together, this thesis shifts the perspective from trying to allocate users' cognitive attention to security and privacy-relevant factors (e.g., Petelka et al., 2019; Caputo et al., 2013) to leveraging socio-emotional factors in influencing the decision-making process.

RQ3: How can we measure employees' information security awareness, taking into account their actual security behavior?

RQ3 was motivated by the current research landscape on security awareness measurement tools, which heavily relies on self-reported assessments and has called for complementary behavior-based approaches (Lebek et al., 2014; Bulgurcu et al., 2010; Kruger and Kearney, 2006). Article D follows this call by presenting and validating a security awareness index that reflects employees' actual security behavior through their susceptibility to simulated social engineering attacks. This contributes to closing the intention-behavior gap criticized by prior works, and provides a less intrusive assessment tool compared to self-reported surveys. Furthermore, the novel methodological approach taken in my thesis confirms prior evaluations of the effectiveness of SETA programs (e.g., Haeussinger and Kranz, 2013; Bulgurcu et al., 2010). Taken together, this thesis uncovers new ways for IS research to sustain the increasing pressure to provide organizations with tools to achieve and maintain information security.

RQ4: How does organizational information security culture respond to external socio-emotional disruptions, such as the COVID-19 pandemic?

Prior research has investigated the impact of internal change, such as the implementation of SETA programs or organizational mergers and acquisitions, on information security culture (Dhillon et al., 2016; Haeussinger and Kranz, 2017). However, the impact of external forces, particularly considering their direction, strength and nature, is hitherto underresearched (Da Veiga et al., 2020). Crucially motivated by the emergence of the COVID-19 pandemic, article E focuses on how global socio-emotional disruption affects information security culture. In doing so, this thesis provides important insights on information security culture as a whole. In particular, it extends prior frameworks by shedding light on novel facilitators and inhibitors of information security culture that emerge in the face of disruption. Zooming into the first year of the pandemic, this thesis unravels how information security culture has been predominantly negatively impacted by disruption in the short term, but – counter-intuitively – can thrive on disruption in the long term. Specifically, the results indicate that disruption can effectively activate crucial security culture components, including top management commitment and cooperation among departments. This is significant as it provides new methods for promoting change in these traditionally resistant areas. Interestingly, the long-term effects of disruption on information security culture are heavily influenced by factors such as the organization's digital maturity and economic stability. These insights are important as they highlight prerequisites needed for a resilient information security culture. In conclusion, this thesis shifts the incumbent

view that information security culture is solely cultivated internally, arguing that it is also shaped by external socio-emotional forces.

7.2 Practical Implications

Beyond the outlined contributions to research, this thesis offers recommendations for information security and privacy practitioners on how to benefit from incorporating socio-emotional factors in their perspective. In particular, it provides valuable insights for (1) designers that aim to support decision-making related to information security and privacy, (2) information security officers, and (3) policymakers and regulators.

This thesis provides several recommendations for designers seeking to assist users in their decision-making through tools (e.g., email reporting tools or password managers) or interventions (e.g., warning about potentially dangerous content or privacy risks). First, the results from studies B and C reveal that, while current tools and interventions primarily focus on influencing cognitive processes, addressing socio-emotional factors might be even more important when trying to assist users in making more secure and privacy-aware choices. For example, the findings from study C indicate that the impact of underlying hedonic motivation is stronger compared to that of utilitarian motivation. In addition, study B demonstrates that users often have a limited understanding of the social ramifications of their decisions to share information, such as the potential violation of their peers' privacy rights. To empower users to make informed decisions, designers of user interfaces should finetune privacy-enhancing interventions to educate users about interdependent privacy issues. Second, study A emphasizes the need for tools and interventions that focus on causing minimum friction, as current approaches tend to intrusively interfere with users' primary goals, such as browsing a website or reading an email. Designers should strive to embed security and privacy-enhancing stimuli seamlessly into users' everyday online activities in a way that requires minimal time and effort to maximize the adoption of these tools. Third, there is a lack of interventions that provide long-term guidance for users in their daily security and privacy decisions. This results in a fragmented landscape of interventions that leaves the user overwhelmed. Instead, designers should adopt a user-centered approach and create tools that guide users throughout their entire security and privacy journey. For example, interventions could be aligned with previously learned educational information, and implications of disclosure decisions could be visualized (e.g., which data is shared with which party after the disclosure) and provided to the user as feedback.

Furthermore, this thesis provides important insights for information security officers. First, the measurement instrument developed in article D provides a valuable tool for evaluating security awareness within an organization. This is a crucial development as it enables organizations to monitor security awareness in real time, while minimizing the intrusiveness, time, and effort required for employees compared to self-reported surveys. Second, the insights from study C underscore the importance of appealing to employees' self-esteem to promote extra-role behavior such as incident reporting. Information security officers can leverage this knowledge by redesigning processes and tools to give users a sense of "warm glow". Third, the findings from study E reveal that external factors such as socio-emotional disruption shape information security culture at large, highlighting the need for information security practitioners to consider the broader social and emotional context. Additionally, the findings indicate that traditionally passive components of information security culture, such as top management commitment and interdepartmental collaboration, can be activated when the stakes are high. This is promising news for practitioners seeking to garner support for security their efforts.

Finally, this thesis addresses regulators and policymakers involved in issues related to information privacy. Currently, regulations do not provide effective solutions to empower individuals to take control of what data about them is collected and stored, what inferences are drawn from it, and how others use it (Spiekermann et al., 2022). Study B addresses this concern and illustrates a significant flaw in current regulations, such as the GDPR, when it comes to privacy violations that are interdependent. In the experimental setup, nearly one fifth of participants chose to disclose their address book, including all of their contacts' phone numbers and other personal information, to Instagram when prompted to do so. This reveals how companies circumvent individuals' privacy rights by collecting users' personal information through the disclosure decisions of their peers. Privacy regulators need to expand their perspective from dyadic information disclosures between individual users and organizations to acknowledging the interconnectedness of sophisticated digital environments and its implications for privacy issues. For example, they should consider requiring two-step opt-in consent for disclosure decisions involving others' personal information, as it has already been implemented for marketing newsletters in the individual privacy context. Recognizing privacy as an interdependent phenomenon becomes particularly crucial when looking at extreme cases of interdependent privacy violations such as Doxing, that is, the disclosure of others' personal information with malicious intent (Douglas, 2016). Establishing a legal foundation for these infringements is necessary to effectively protect individuals' privacy rights.

7.3 Limitations and Future Research

Although this thesis provides valuable theoretical and practical contributions, the results should be interpreted in light of their limitations. Accordingly, three noteworthy limitations and corresponding avenues for future research are outlined next.

First, the studies incorporated in this thesis are subject to methodological limitations. In studies B and C, the respective online vignette experimental setup captured users' responses to the presented experimental context in one-time interactions. For example, users were prompted to synchronize their address book with Instagram only once. While the results of these studies make an important first step in characterizing users' behavior, only longitudinal investigations can confirm whether the observed effects from one-time interactions result in sustainable changes in behavior over time (Karahanna et al., 2018a). For example, future research could employ experience sampling approaches to capture users' security and privacy behavior throughout their daily life (Benlian, 2020; Benlian, 2022). In the organizational context, studies D and E took a longitudinal approach to investigate information security awareness and culture across multiple organizations. While these studies provide valuable insights into the dynamics of the human factor in organizational information security, they narrowed their focus to the perspective of information security officers. To broaden our understanding of how organizational information security responds to disruption, we invite researchers to take these findings as a basis to conduct further research using methods that simultaneously consider multiple views, such as the management, IT function, and employee perspective (Benlian, 2013; Benlian and Haffke, 2016). Additionally, the studies in this thesis recruited participants from Western societies (e.g., USA, EU). Given that security and privacy behavior can be subject to cultural influence (Krasnova et al., 2012), future research should consider studying different cultural backgrounds to verify the generalizability of the results of this thesis.

Second, this thesis only begins to unravel the underlying mechanisms that dictate and explain information security and privacy behavior driven by socio-emotional influences. While insights into the mediating effects of hedonic vs. utilitarian motivation and the salience of others within one's socio-technical environment contribute to a more holistic understanding of security and privacy, future research can further expand our body of knowledge on additional moderating factors. For example, how users respond to socio-emotional triggers may be subject to their personality. Moreover, group-level conditions such as social norms may play an important role and influence users' behavior differently when different social identities are activated (Bélanger and James, 2020). By delving deeper into socio-emotional circumstances and investigating

influential boundary conditions, we could gain a more holistic understanding of what promotes and what restrains information security and privacy.

Finally, the studies on security and privacy decision-making presented in this thesis narrow their scope to investigating how socio-emotional factors affect specifically users' information disclosure and protection in two selected contexts. To broaden our understanding of the implications of the human aspect in security and privacy, future research could go beyond user-specific outcomes and examine, for example, strategies for online service providers to intentionally implement security and privacy-enhancing features with the aim to increase user engagement through user empowerment (Werner et al., 2022). Such strategies could help make security and privacy-enhancing features attractive even to providers that traditionally prioritize the collection of personal user data.

7.4 Conclusion

This thesis is one of the first attempts to systematically investigate the importance and ramifications of socio-emotional factors in information security and privacy. The first strand of this thesis provides insights on how social and emotional factors (e.g., a sense of "warm glow" after performing an altruistic act) influence users' decision-making related to security and privacy. The results demonstrate that when tools and interventions are designed with the aim to specifically leverage these factors, users are not only more cautious in disclosing information, but also more likely to proactively promote information security and privacy within their organization. These findings provide important insights for sustainably mitigating information security and privacy risks that current technology and legislation cannot control. The second strand of this thesis uncovers, first, the potential of social engineering simulations as a suitable and nonintrusive means to measure organizational security awareness, which enables organizations to continuously manage the human dimension of their information security risk landscape. Second, it sheds light on how socio-emotional disruptions (e.g., the COVID-19 pandemic) impact organizational information security culture. Counter-intuitively, the results highlight that the activation of crucial security culture components through disruption can yield long-term improvement if certain prerequisites are met. These findings provide valuable insights for information security practitioners seeking to build a cyber-resilient organization. In conclusion, this thesis aims to inspire future research to unlock the untapped potential of social and emotional factors in protecting the information assets that form the foundation for our digital world.

References

Abbasi, A., Dobolyi, D., Vance, A. and Zahedi, F. M. (2021). "The Phishing Funnel Model: A Design Artifact to Predict User Susceptibility to Phishing Websites", *Information Systems Research* 32 (2)**,** 410-436.

Abbasi, A., Zahedi, F. M. and Chen, Y. (2016). "Phishing Susceptibility: The Good, the Bad, and the Ugly", *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, Seattle, WA, USA.

Abomhara, M. and Køien, G. M. (2015). "Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attacks", *Journal of Cyber Security and Mobility* 4 (1)**,** 65–88.

Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F. and Sleeper, M. (2017). "Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online", *ACM Computing Surveys (CSUR)* 50 (3)**,** 1-41.

Adam, M., Werner, D., Wendt, C. and Benlian, A. (2020). "Containing Covid-19 through Physical Distancing: The Impact of Real-Time Crowding Information", *European Journal of Information Systems* 29 (5)**,** 595-607.

Agarwal, R. and Karahanna, E. (2000). "Time Flies When You're Having Fun: Cognitive Absorption and Beliefs About Information Technology Usage", *MIS Quarterly* 24 (4)**,** 665-694.

Ågerfalk, P. J., Conboy, K. and Myers, M. D. (2020). "Information Systems in the Age of Pandemics: Covid-19 and Beyond", *European Journal of Information Systems* 29 (3).

Aguinis, H. and Bradley, K. J. (2014). "Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies", *Organizational Research Methods* 17 (4)**,** 351-371.

Algarni, A., Xu, Y. and Chan, T. (2017). "An Empirical Study on the Susceptibility to Social Engineering in Social Networking Sites: The Case of Facebook", *European Journal of Information Systems* 26 (6)**,** 661-687.

AlHogail, A. (2015). "Design and Validation of Information Security Culture Framework", *Computers in Human Behavior* 49**,** 567-575.

Allodi, L., Chotza, T., Panina, E. and Zannone, N. (2019). "The Need for New Antiphishing Measures against Spear-Phishing Attacks", *IEEE Security & Privacy* 18 (2)**,** 23-34.

Almuhimedi, H., Schaub, F., Sadeh, N., Adjerid, I., Acquisti, A., Gluck, J., Cranor, L. F. and Agarwal, Y. (2015). "Your Location Has Been Shared 5,398 Times! A Field Study on Mobile App Privacy Nudging", *33rd ACM Conference on Human Factors in Computing Systems*, Seoul, South Korea.

Alnajim, A. and Munro, M. (2009). "An Anti-Phishing Approach That Uses Training Intervention for Phishing Websites Detection", *Sixth International Conference on Information Technology: New Generations*, Las Vegas, NV, USA.

Alsarkal, Y., Zhang, N. and Xu, H. (2018). "Your Privacy Is Your Friend's Privacy: Examining Interdependent Information Disclosure on Online Social Networks", *51st Hawaii International Conference on System Sciences*, Hawaii, USA.

Alshaikh, M. (2020). "Developing Cybersecurity Culture to Influence Employee Behavior: A Practice Perspective", *Computers & Security* 98**,** 102003.

Andreoni, J. (1990). "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving", *The Economic Journal* 100 (401)**,** 464-477.

APWG. (2019). *Phishing Activity Trends Report, 3rd Quarter 2019*. [Online] Available: https://docs.apwg.org/reports/apwg_trends_report_q3_2019.pdf (Accessed January 19, 2020).

APWG. (2020). *Phishing Activity Trends Report, 3rd Quarter 2020*. [Online] Available: https://docs.apwg.org/reports/apwg_trends_report_q3_2020.pdf (Accessed January 11, 2023).

APWG. (2022). *Phishing Activity Trends Report, 3rd Quarter 2022*. [Online] Available: https://docs.apwg.org/reports/apwg_trends_report_q3_2022.pdf (Accessed January 11, 2023).

Arachchilage, N. A. G. and Love, S. (2013). "A Game Design Framework for Avoiding Phishing Attacks", *Computers in Human Behavior* 29 (3)**,** 706-714.

Arachchilage, N. A. G., Love, S. and Beznosov, K. (2016). "Phishing Threat Avoidance Behaviour: An Empirical Investigation", *Computers in Human Behavior* 60**,** 185-197.

Avgerou, C. (2013). "Social Mechanisms for Causal Explanation in Social Theory Based IS Research", *Journal of the Association for Information Systems* 14 (8)**,** 3.

Barev, T. J., Janson, A. and Leimeister, J. M. (2020). "Designing Effective Privacy Nudges in Digital Environments: A Design Science Research Approach", *International Conference on Design Science Research in Information Systems and Technology*, Kristiansand/Virtual.

Baslyman, M. and Chiasson, S. (2016). "Smells Phishy?: An Educational Game About Online Phishing Scams", *APWG Symposium on Electronic Crime Research (eCrime)*, Toronto, ON, Canada.

Bedford, J., Enria, D., Giesecke, J., Heymann, D. L., Ihekweazu, C., Kobinger, G., Lane, H. C., Memish, Z., Oh, M.-d. and Schuchat, A. (2020). "Covid-19: Towards Controlling of a Pandemic", *The Lancet* 395**,** 1015-1018.

Beguin, E., Besnard, S., Cros, A., Joannes, B., Leclerc-Istria, O., Noel, A., Roels, N., Taleb, F., Thongphan, J., Alata, E. and others (2019). "Computer-Security-Oriented Escape Room", *IEEE Security & Privacy* 17 (4)**,** 78-83.

Belanger, F., Hiller, J. S. and Smith, W. J. (2002). "Trustworthiness in Electronic Commerce: The Role of Privacy, Security, and Site Attributes", *The journal of strategic Information Systems* 11 (3-4)**,** 245-270.

Bélanger, F. and James, T. L. (2020). "A Theory of Multilevel Information Privacy Management for the Digital Era", *Information systems research* 31 (2)**,** 510-536.

Bélanger, F., Maier, J. and Maier, M. (2022). "A Longitudinal Study on Improving Employee Information Protective Knowledge and Behaviors", *Computers & Security* 116**,** 102641.

Benlian, A. (2013). "Effect Mechanisms of Perceptual Congruence between Information Systems Professionals and Users on Satisfaction with Service", *Journal of Management Information Systems* 29 (4)**,** 63-96.

Benlian, A. (2020). "A Daily Field Investigation of Technology-Driven Spillovers from Work to Home", *MIS Quarterly* 44 (3).

Benlian, A. (2022). "Sprint Zeal or Sprint Fatigue? The Benefits and Burdens of Agile Isd Practices Use for Developer Well-Being", *Information Systems Research* 33 (2)**,** 557-578.

Benlian, A. and Haffke, I. (2016). "Does Mutuality Matter? Examining the Bilateral Nature and Effects of CEO–Cio Mutual Understanding", *The Journal of Strategic Information Systems* 25 (2)**,** 104-126.

Benlian, A., Kettinger, W. J., Sunyaev, A., Winkler, T. J. and Editors, G. (2018). "The Transformative Value of Cloud Computing: A Decoupling, Platformization, and Recombination Theoretical Framework", *Journal of Management Information Systems* 35 (3)**,** 719-739.

Benlian, A., Klumpe, J. and Hinz, O. (2020). "Mitigating the Intrusive Effects of Smart Home Assistants by Using Anthropomorphic Design Features: A Multimethod Investigation", *Information Systems Journal* 30 (6)**,** 1010-1042.

Benlian, A., Wiener, M., Cram, W. A., Krasnova, H., Maedche, A., Mohlmann, M., Recker, J. and Remus, U. (2022). "Algorithmic Management: Bright and Dark Sides, Practical Implications, and Research Opportunities", *Business & Information Systems Engineering* 64 (6)**,** 825-839.

Bennett, C. J. (2016). "Voter Databases, Micro-Targeting, and Data Protection Law: Can Political Parties Campaign in Europe as They Do in North America?", *International Data Privacy Law* 6 (4)**,** 261-275.

Biczók, G. and Chia, P. H. (2013). "Interdependent Privacy: Let Me Share Your Data", *International Conference on Financial Cryptography and Data Security*, Okinawa, Japan.

Biczók, G., Huguenin, K., Humbert, M. and Grossklags, J. (2021). "Call for Papers: Special Issue on Managing Multi-Party, Interdependent Privacy Risks", *Computers & Security*.

Biselli, T. and Reuter, C. (2021). "On the Relationship between It Privacy and Security Behavior: A Survey among German Private Users", *16. Internationale Tagung Wirtschaftsinformatik*, Duisburg / Essen, Deutschland.

Bitkom. (2018). *Spionage, Sabotage Und Datendiebstahl – Wirtschaftsschutz in Der Industrie. Studienbericht 2018, Bitkom E.V.* . [Online] Available: https://www.bitkom.org/sites/default/files/file/import/181008-Bitkom-Studie-Wirtschaftsschutz-2018-NEU.pdf (Accessed January 01, 2020).

Blythe, J., Camp, J. and Garg, V. (2011). "Targeted Risk Communication for Computer Security", *16th International Conference on Intelligent User Interfaces,* Palo Alto, CA, USA.

Briggs, P., Jeske, D. and Coventry, L. (2017). "The Design of Messages to Improve Cybersecurity Incident Reporting", *International Conference on Human Aspects of Information Security, Privacy, and Trust,* Vancouver, BC, Canada.

Brocke, J. v., Simons, A., Niehaves, B., Reimer, K., Plattfaut, R. and Cleven, A. (2009). "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process", *European Conference on Information Systems,* Verona, Italy.

BSI. (2019). *Pressemitteilung*. [Online] Available: https://www.bsi.bund.de/DE/Presse/Pressemitteilungen/Presse2019/Spam-Bundesbehoerden_181219.html (Accessed January 13, 2020).

Buckman, J. R., Bockstedt, J. C. and Hashim, M. J. (2019). "Relative Privacy Valuations under Varying Disclosure Characteristics", *Information Systems Research* 30 (2)**,** 375-388.

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

Burda, P., Allodi, L. and Zannone, N. (2020). "Don't Forget the Human: A Crowdsourced Approach to Automate Response and Containment against Spear Phishing Attacks", *IEEE European Symposium on Security and Privacy Workshops (EuroS&PW),* Virtual.

Burns, A., Roberts, T. L., Posey, C. and Lowry, P. B. (2019a). "The Adaptive Roles of Positive and Negative Emotions in Organizational Insiders' Security-Based Precaution Taking", *Information Systems Research* 30 (4)**,** 1228-1247.

Burns, A. J., Johnson, M. E. and Caputo, D. D. (2019b). "Spear Phishing in a Barrel: Insights from a Targeted Phishing Campaign", *Journal of Organizational Computing and Electronic Commerce* 29 (1)**,** 24-39.

Burns, M. B., Durcikova, A. and Jenkins, J. L. (2012). "On Not Falling for Phish: Examining Multiple Stages of Protective Behavior of Information Systems End-Users", *33rd International Conference on Information Systems,* Orlando, FL, USA.

Burns, M. B., Durcikova, A. and Jenkins, J. L. (2013). "What Kind of Interventions Can Help Users from Falling for Phishing at Tempts: A Research Proposal for Examining Stage-Appropriate Interventions", *46th Hawaii International Conference on System Sciences,* Wailea, HI, USA.

Canova, G., Volkamer, M., Bergmann, C. and Reinheimer, B. (2015). "NoPhish App Evaluation: Lab and Retention Study", *NDSS Workshop on Usable Security,* San Diego, CA, USA.

Cao, Z., Hui, K.-L. and Xu, H. (2018). "An Economic Analysis of Peer Disclosure in Online Social Communities", *Information Systems Research* 29 (3)**,** 546-566.

Caputo, D. D., Pfleeger, S. L., Freeman, J. D. and Johnson, M. E. (2013). "Going Spear Phishing: Exploring Embedded Training and Awareness", *IEEE Security & Privacy*, San Francisco, USA.

Caraban, A., Karapanos, E., Gonçalves, D. and Campos, P. (2019). "23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction", *CHI Conference on Human Factors in Computing Systems*, Glasgow, UK.

Carella, A., Kotsoev, M. and Truta, T. M. (2017). "Impact of Security Awareness Training on Phishing Click-through Rates", *IEEE International Conference on Big Data*, Boston, MA, USA.

Carroll, N. and Conboy, K. (2020). "Normalising the "New Normal": Changing Tech-Driven Work Practices under Pandemic Time Pressure", *International Journal of Information Management***,** 102186.

Casciano, R. and Massey, D. S. (2012). "Neighborhood Disorder and Anxiety Symptoms: New Evidence from a Quasi-Experimental Study", *Health & Place* 18 (2)**,** 180-190.

Chen, H. and Li, W. (2019). "Understanding Commitment and Apathy in Is Security Extra-Role Behavior from a Person-Organization Fit Perspective", *Behaviour & Information Technology* 38 (5)**,** 454-468.

Chen, J., Ping, J. W., Xu, Y. and Tan, B. C. (2015a). "Information Privacy Concern About Peer Disclosure in Online Social Networks", *IEEE Transactions on Engineering Management* 62 (3)**,** 311-324.

Chen, Y., Ramamurthy, K. and Wen, K.-W. (2015b). "Impacts of Comprehensive Information Security Programs on Information Security Culture", *Journal of Computer Information Systems* 55 (3)**,** 11-19.

Choe, E. K., Jung, J., Lee, B. and Fisher, K. (2013). "Nudging People Away from Privacy-Invasive Mobile Apps through Visual Framing", *IFIP Conference on Human-Computer Interaction*, Cape Town, South Africa.

Cloudmark. (2016). *Spear Phishing: The Secret Weapon Behind the Worst Cyber Attacks*. [Online] Available: https://www.cloudmark.com/en/blog/spear-phishing-secret-weapon-behind-worst-cyber-attacks (Accessed January 13, 2020).

Cofense. (2020). *Cofense Q3 Phishing Review*. https://go.cofense.com/wp-content/uploads/pdf/Cofense-Q3_2020_Phishing -Review-report.pdf. [Online] Available: https://go.cofense.com/wp-content/uploads/pdf/Cofense-Q3_2020_Phishing-Review-report.pdf (Accessed February 15, 2021).

Conboy, K. (2019). "Being Promethean", *European Journal of Information Systems* 28 (2)**,** 119-125.

Corbin, J. and Strauss, A. (2014). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*: SAGE Publications.

Coventry, L., Briggs, P., Jeske, D. and Moorsel, A. (2014). "SCENE: A Structured Means for Creating and Evaluating Behavioral Nudges in a Cyber Security Environment", *International Conference of Design, User experience, and Usability*, Heraklion, Greece.

Craigen, D., Diakun-Thibault, N. and Purse, R. (2014). "Defining Cybersecurity", *Technology Innovation Management Review* 4 (10).

Cram, A., Proudfoot, J. G. and Bentley, U. (2019). "Seeing the Forest and the Trees: A Meta-Analysis of the Antecedents to Information Security Policy Compliance", *MIS Quarterly* 43 (2)**,** 525-554.

Cram, W. A., Proudfoot, J. G. and D'arcy, J. (2017). "Organizational Information Security Policies: A Review and Research Framework", *European Journal of Information Systems* 26 (6)**,** 605-641.

Cranor, L. F. and Garfinkel, S. (2005). *Security and Usability: Designing Secure Systems That People Can Use*: O'Reilly Media, Inc.

Cronbach, L. J. (1951). "Coefficient Alpha and the Internal Structure of Tests", *Psychometrika* 16 (3)**,** 297-334.

Cuchta, T., Blackwood, B., Devine, T. R., Niichel, R. J., Daniels, K. M., Lutjens, C. H., Maibach, S. and Stephenson, R. J. (2019). "Human Risk Factors in Cybersecurity", *20th Annual SIG Conference on Information Technology Education*, Tacoma, WA, USA.

Da Veiga, A., Astakhova, L. V., Botha, A. and Herselman, M. (2020). "Defining Organisational Information Security Culture—Perspectives from Academia and Industry", *Computers & Security* 92**,** 101713.

Da Veiga, A. and Eloff, J. H. P. (2010). "A Framework and Assessment Instrument for Information Security Culture", *Computers & Security* 29 (2)**,** 196-207.

Dalal, R. S., Howard, D. J., Bennett, R. J., Posey, C., Zaccaro, S. J. and Brummel, B. J. (2022). "Organizational Science and Cybersecurity: Abundant Opportunities for Research at the Interface", *Journal of Business and Psychology* 37 (1)**,** 1-29.

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

Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology", *MIS Quarterly* 13 (3)**,** 319-340.

De Ryck, P., Nikiforakis, N., Desmet, L. and Joosen, W. (2013). "Tabshots: Client-Side Detection of Tabnabbing Attacks", *8th ACM SIGSAC Symposium on Information, Computer and Communications Security*, Hangzhou, China.

Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F. and Holz, T. (2019). "We Value Your Privacy... Now Take Some Cookies-Measuring the Gdpr's Impact on Web Privacy", *Informatik Spektrum* 42 (5).

Dennis, A. R. and Minas, R. K. (2018). "Security on Autopilot: Why Current Security Theories Hijack Our Thinking and Lead Us Astray", *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* 49 (SI)**,** 15-38.

Desouza, K. C., Ahmad, A., Naseer, H. and Sharma, M. (2020). "Weaponizing Information Systems for Political Disruption: The Actor, Lever, Effects, and Response Taxonomy (Alert)", *Computers & Security* 88**,** 101606.

Dhamija, R. and Tygar, J. D. (2005). "The Battle against Phishing: Dynamic Security Skins", *Symposium on Usable Privacy and Security*, Pittsburgh, PA, USA.

Dhillon, G., Syed, R. and Pedron, C. (2016). "Interpreting Information Security Culture: An Organizational Transformation Case Study", *Computers & Security* 56**,** 63-69.

Dickinger, A., Arami, M. and Meyer, D. (2008). "The Role of Perceived Enjoyment and Social Norm in the Adoption of Technology with Network Externalities", *European Journal of Information Systems* 17 (1)**,** 4-11.

Dickinson, D. A. (2020). "The Role of Enterprise Social Networks (Esn) in Maintaining Organizational Rhythms During the Covid-19 Pandemic", *Academy of Management - Rapid Research Plenary: COVID 19 and Organizational Behavior*.

Dienlin, T. and Metzger, M. J. (2016). "An Extended Privacy Calculus Model for Snss: Analyzing Self-Disclosure and Self-Withdrawal in a Representative Us Sample", *Journal of Computer-Mediated Communication* 21 (5)**,** 368-383.

Dinev, T. and Hart, P. (2006). "An Extended Privacy Calculus Model for E-Commerce Transactions", *Information Systems Research* 17 (1)**,** 61-80.

Dinev, T., McConnell, A. R. and Smith, H. J. (2015). "Research Commentary—Informing Privacy Research through Information Systems, Psychology, and Behavioral Economics: Thinking Outside the "Apco" Box", *Information Systems Research* 26 (4)**,** 639-655.

Douglas, D. M. (2016). "Doxing: A Conceptual Analysis", *Ethics and Information Technology* 18 (3)**,** 199-210.

Egelman, S., Cranor, L. F. and Hong, J. (2008). "You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings", *SIGCHI Conference on Human Factors in Computing Systems*, Florence, Italy.

Egelman, S. and Peer, E. (2015). "The Myth of the Average User: Improving Privacy and Security Systems T Hrough Individualization", *New Security Paradigms Workshop*, Twente, The Netherlands.

Ellison, N. B., Vitak, J., Steinfield, C., Gray, R. and Lampe, C. (2011). "Negotiating Privacy Concerns and Social Capital Needs in a Social Media Environment", (eds.) *Privacy Online*, 19-32. Springer.

ENISA. (2022). *Enisa Threat Landscape 2022*. [Online] Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022 (Accessed January 15, 2023).

Esteve, A. (2017). “The Business of Personal Data: Google, Facebook, and Privacy Issues in the EU and the USA”, *International Data Privacy Law* 7 (1)**,** 36-47.

EUR-Lex. (2012). *Charter of Fundamental Rights of the European Union*. [Online] Available: https://eur-lex.europa.eu/eli/treaty/char_2012/oj (Accessed July 19, 2021).

European Parliament and Council (2016). “Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/Ec (General Data Protection Regulation)”, *Official Journal of the European Union*.

Fatima, R., Yasin, A., Liu, L. and Wang, J. (2019). “How Persuasive Is a Phishing Email? A Phishing Game for Phishing Aware Ness”, *Journal of Computer Security* 27 (6)**,** 581-612.

Fishbein, M. and Ajzen, I. (1977). “Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research”, *Philosophy and Rhetoric* 10 (2).

Foss, N. J. (2020). “The Impact of the Covid-19 Pandemic on Firms’ Organizational Designs”, *Journal of Management Studies* 58 (1)**,** 270-274.

Franz, A. (2019). “Security Awareness Messbar Machen - Der Employee Security Index”, *Tagungsband des 16. Deutschen IT-Sicherheitskongress*, Bonn, Deutschland.

Franz, A., Zimmermann, V., Albrecht, G., Hartwig, K., Reuter, C., Benlian, A. and Vogt, J. (2021). “Sok: Still Plenty of Phish in the Sea—a Taxonomy of User-Oriented Phishing Interventions and Avenues for Future Research”, *Seventeenth Symposium on Usable Privacy and Security*, Virtual.

Gal-Or, E., Gal-Or, R. and Penmetsa, N. (2018). “The Role of User Privacy Concerns in Shaping Competition among Platforms”, *Information Systems Research* 29 (3)**,** 698-722.

Garcia, D. (2017). “Leaking Privacy and Shadow Profiles in Online Social Networks”, *Science Advances* 3 (8)**,** e1701172.

Gastellier-Prevost, S., Granadillo, G. G. and Laurent, M. (2011). “A Dual Approach to Detect Pharming Attacks at the Client-Side”, *4th IFIP International Conference on New Technologies, Mobility and Security*, Paris, France.

Gaver, W. W. (1991). “Technology Affordances”, *Proceedings of the SIGCHI conference on Human factors in computing systems*, New Orleans, USA.

Gibson, C. (2020). “From “Social Distancing” to “Care in Connecting”: An Emerging Organizational Research Agenda for Turbulent Times”, *Academy of Management Discoveries* 6 (2)**,** 165-169.

Gibson, J. J. (1979). “The Theory of Affordances. The Ecological Approach to Visual Perception”, (eds.) *The People, Place and, Space Reader*, 56-60. Routledge New York and London.

Gleasure, R. and Feller, J. (2016). “Does Heart or Head Rule Donor Behaviors in Charitable Crowdfunding Markets?”, *International Journal of Electronic Commerce* 20 (4)**,** 499-524.

Goel, S., Williams, K., University at Albany, S., Dincelli, E. and University at Albany, S. (2017). "Got Phished? Internet Security and Human Vulnerability", *Journal of the Association for Information Systems* 18 (1)**,** 22-44.

Gokul, C., Pandit, S., Vaddepalli, S., Tupsamudre, H., Banahatti, V. and Lodha, S. (2018). "Phishy - a Serious Game to Train Enterprise Users on Phishing Awareness", *Annual Symposium on Computer-human Interaction in Play (CHI PLAY)*, Melbourne, Australia.

Greene, K., Steves, M. and Theofanos, M. (2018a). "No Phishing Beyond This Point", *Computer* 51 (6)**,** 86-89.

Greene, K. K., Steves, M. P., Theofanos, M. F. and Kostick, J. (2018b). "User Context: An Explanatory Variable in Phishing Susceptibility", *Workshop Usable Security.*, London, England.

Gregory, R. W., Kaganer, E., Henfridsson, O. and Ruch, T. J. (2018). "It Consumerization and the Transformation of It Governance", *MIS Quarterly* 42 (4)**,** 1225-1253.

Guillemette, M. and Pare, G. (2005). "Understanding the Role and Transformation of the Information Technology Function in Organizations", *International Conference On Information Systems*, Las Vegas, USA.

Haeussinger, F. and Kranz, J. (2013). "Information Security Awareness: Its Antecedents and Mediating Effects on Security Compliant Behavior", *34th International Conference on Information Systems*, Milan, Italy.

Haeussinger, F. and Kranz, J. (2017). "Antecedents of Employees' Information Security Awareness - Review, Synthesis, and Directions for Future Research", *European Conference On Information Systems*, Guimaraes, Portugal.

Hale, M. and Gamble, R. (2014). "Toward Increasing Awareness of Suspicious Content through Game Play", *IEEE World Congress on Services*, Anchorage, AK, USA.

Hale, M. L., Gamble, R. F. and Gamble, P. (2015). "Cyberphishing: A Game-Based Platform for Phishing Awareness Testing", *48th Hawaii International Conference on System Sciences*, Kauai, HI, USA.

Harwell, D. and Harris, S. (2021). *White House Has Spoken to Israeli Officials About Spyware Concerns Following Pegasus Project Revelations*. The Washington Post. [Online] Available: https://www.washingtonpost.com/technology/2021/07/29/pegasus-white-house-israel-concerns/ (Accessed 06.08.2021).

Hassandoust, F., Techatassanasoontorn, A. A. and Singh, H. (2020). "Information Security Behaviour: A Critical Review and Research Directions", *European Conference on Information Systems*, Virtual conference.

Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*: Guilford Publications.

Heartfield, R. and Loukas, G. (2018). "Detecting Semantic Social Engineering Attacks with the Weakest Link: Implementation and Empirical Evaluation of a Human-as-a-Security-Sensor Framework", *Computers & Security* 76**,** 101-127.

Heise. (2019a). *Computervirus Klinikum Fürth Offline Und Mit Eingeschränktem Betrieb*. [Online] Available: https://www.heise.de/newsticker/meldung/Computervirus-Klinikum-Fuerth-offline-und-mit-eingeschraenktem-Betrieb-4615427.html (Accessed January 13, 2020).

Heise. (2019b). *It-Systeme Der Stadt Frankfurt Am Main Wegen Malware-Befall Offline*. [Online] Available: https://www.heise.de/newsticker/meldung/IT-Systeme-der-Stadt-Frankfurt-am-Main-wegen-Malware-Befall-offline-4619634.html (Accessed January 13, 2020).

Heise. (2019c). *Uni Gießen Nähert Sich Nach Hacker-Attacke Wieder Dem Normalbetrieb*. [Online] Available: https://www.heise.de/newsticker/meldung/Uni-Giessen-naehert-sich-nach-Hacker-Attacke-wieder-dem-Normalbetrieb-4628715.html (Accessed January 13, 2020).

Herath, T. and Rao, H. R. (2009). "Protection Motivation and Deterrence: A Framework for Security Policy Compliance in Organisations", *European Journal of Information Systems* 18 (2)**,** 106-125.

Hertel, M. (2017). "Risiken Der Industrie 4.0–Eine Strukturierung Von Bedrohungsszenarien Der Smart Factory", (eds.) *It-Grc-Management–Governance, Risk Und Compliance*, 113-128. Springer.

Herzberg, A. and Jbara, A. (2008). "Security and Identification Indicators for Browsers against Spoofing a Nd Phishing Attacks", *ACM Transactions on Internet Technology (TOIT)* 8 (4)**,** 1-36.

Herzberg, A. and Margulies, R. (2013). "Forcing Johnny to Login Safely", *Journal of Computer Security* 21 (3)**,** 393-424.

Hsu, J. S. C., Shih, S. P., Hung, Y. W. and Lowry, P. B. (2015). "The Role of Extra-Role Behaviors and Social Controls in Information Security Policy Effectiveness", *Information Systems Research* 26 (2)**,** 282-300.

Hu, P. J., Chau, P. Y., Sheng, O. R. L. and Tam, K. Y. (1999). "Examining the Technology Acceptance Model Using Physician Acceptance of Telemedicine Technology", *Journal of Management Information Systems* 16 (2)**,** 91-112.

Hu, Q., Dinev, T., Hart, P. and Cooke, D. (2012). "Managing Employee Compliance with Information Security Policies: The Critical Role of Top Management and Organizational Culture", *Decision Sciences* 43 (4)**,** 615-660.

Huang, K. and Pearlson, K. (2019). "For What Technology Can't Fix: Building a Model of Organizational Cybersecurity Culture", *Hawaii International Conference on System Sciences*, Maui, USA.

Humbert, M., Trubert, B. and Huguenin, K. (2019). "A Survey on Interdependent Privacy", *ACM Computing Surveys (CSUR)* 52 (6)**,** 1-40.

Iacono, L. L., Nguyen, H. V., Hirsch, T., Baiers, M. and Möller, S. (2014). "UI-Dressing to Detect Phishing", *6th International Symposium on Cyberspace Safety and Security*, Paris, France.

Imprivata. (2020). *How Reputational Damage from a Data Breach Affects Consumer Perception*. [Online] Available: https://www.imprivata.com/blog/reputation-risks-how-cyberattacks-affect-consumer-perception (Accessed January 08, 2023).

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

ISF (2007). Information Security Forum (Isf) Standard of Good Practice 2007, Cb3.4.

ISO. (2013). *Iso/Iec 27001:2013*. [Online] Available: https://www.iso.org/standard/54534.html (Accessed 12.11.2020).

Iweala, S., Spiller, A. and Meyerding, S. (2019). "Buy Good, Feel Good? The Influence of the Warm Glow of Giving on the Evaluation of Food Items with Ethical Claims in the Uk and Germany", *Journal of Cleaner Production* 215**,** 315-328.

Jakobsson, M. and Myers, S. (2007). "Delayed Password Disclosure", *ACM SIGACT News* 38 (3)**,** 56-75.

Jansen, J. and Schaik, P. (2019). "The Design and Evaluation of a Theory-Based Intervention to Promote Se Curity Behaviour against Phishing", *International Journal of Human-Computer Studies* 123**,** 40-55.

Jenkins, J. L., Durcikova, A. and Nunamaker, J. F. (2021). "Mitigating the Security Intention-Behavior Gap: The Moderating Role of Required Effort on the Intention-Behavior Relationship", *Journal of the Association for Information Systems* 22 (1).

Jensen, M., Durcikova, A. and Wright, R. (2017a). "Combating Phishing Attacks: A Knowledge Management Approach", *50th Hawaii International Conference on System Sciences*, HI, USA.

Jensen, M. L., Dinger, M., Wright, R. T. and Thatcher, J. B. (2017b). "Training to Mitigate Phishing Attacks Using Mindfulness Techniques", *Journal of Management Information Systems* 34 (2)**,** 597-626.

Jeske, D., Coventry, L. and Briggs, P. (2014). "Nudging Whom How : It Proficiency , Impulse Control and Secure Behavio Ur",

Jia, H. and Xu, H. (2016). "Autonomous and Interdependent: Collaborative Privacy Management on Social Networking Sites", *CHI Conference on Human Factors in Computing Systems*, San Jose, USA.

Jiang, Z., Heng, C. S. and Choi, B. C. (2013). "Research Note—Privacy Concerns and Privacy-Protective Behavior in Synchronous Online Social Interactions", *Information Systems Research* 24 (3)**,** 579-595.

Johnston, Johnston, A. C., Di Gangi, P. M., University of Alabama at, B., Howard, J., University of Alabama at, B., Worrell, J. and University of Alabama at, B. (2019). "It Takes a Village: Understanding the Collective Security Efficacy of Employee Groups", *Journal of the Association for Information Systems***,** 186-212.

Johnston, A. C., Warkentin, M., McBride, M. and Carter, L. (2016). "Dispositional and Situational Factors: Influences on Information Security Policy Violations", *European Journal of Information Systems* 25 (3)**,** 231-251.

Kahneman, D. (2011). *Thinking, Fast and Slow*: Macmillan.

Kam, H. J., Ormond, D. K., Menard, P. and Crossler, R. E. (2021). "That's Interesting: An Examination of Interest Theory and Self-Determination in Organisational Cybersecurity Training", *Information Systems Journal* 32 (4).

Kamleitner, B. and Mitchell, V. (2019). "Your Data Is My Data: A Framework for Addressing Interdependent Privacy Infringements", *Journal of Public Policy & Marketing* 38 (4)**,** 433-450.

Kamleitner, B. and Sotoudeh, M. (2019). "Information Sharing and Privacy as a Socio-Technical Phenomenon: Interview with Bernadette Kamleitner", *TATuP-Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis* 28 (3)**,** 68-71.

Kankane, S., DiRusso, C. and Buckley, C. (2018). "Can We Nudge Users toward Better Password Management? An Initial Study", *CHI Conference on Human Factors in Computing Systems*, Montréal, Canada.

Karahanna, E., Benbasat, I., Bapna, R. and Rai, A. (2018a). "Editor's Comments: Opportunities and Challenges for Different Types of Online Experiments", *MIS Quarterly* 42 (4)**,** iii-x.

Karahanna, E. and Straub, D. W. (1999). "The Psychological Origins of Perceived Usefulness and Ease-of-Use", *Information & Management* 35 (4)**,** 237-250.

Karahanna, E., Xu, S. X., Xu, Y. and Zhang, N. A. (2018b). "The Needs–Affordances–Features Perspective for the Use of Social Media", *MIS Quarterly* 42 (3)**,** 737-756.

Karjalainen, M., Sarker, S. and Siponen, M. (2019). "Toward a Theory of Information Systems Security Behaviors of Organizational Employees: A Dialectical Process Perspective", *Information Systems Research* 30 (2)**,** 687-704.

Karlsson, F., Åström, J. and Karlsson, M. (2015). "Information Security Culture–State-of-the-Art Review between 2000 and 2013", *Information & Computer Security* 23 (3).

Kehr, F., Kowatsch, T., Wentzel, D. and Fleisch, E. (2015). "Blissfully Ignorant: The Effects of General Privacy Concerns, General Institutional Trust, and Affect in the Privacy Calculus", *Information Systems Journal* 25 (6)**,** 607-635.

Kirlappos, I., Beautement, A. and Sasse, M. A. (2013). ""Comply or Die" Is Dead: Long Live Security-Aware Principal Agents", *International Conference on Financial Cryptography and Data Security*, Okinawa, Japan.

Kirlappos, I. and Sasse, M. A. (2011). "Security Education against Phishing: A Modest Proposal for a Major Rethink", *IEEE Security & Privacy* 10 (2)**,** 24-32.

Knijnenburg, B. P., Page, X., Wisniewski, P., Lipford, H. R., Proferes, N. and Romano, J. (2022). *Modern Socio-Technical Perspectives on Privacy*: Springer Nature.

Krasnova, H., Veltri, N. F. and Günther, O. (2012). "Self-Disclosure and Privacy Calculus on Social Networking Sites: The Role of Culture", *Business & Information Systems Engineering* 4 (3)**,** 127-135.

Kroll, T. and Stieglitz, S. (2021). "Digital Nudging and Privacy: Improving Decisions About Self-Disclosure in Social Networks", *Behaviour & Information Technology* 40 (1)**,** 1-19.

Kruger, H. A. and Kearney, W. D. (2006). "A Prototype for Assessing Information Security Awareness", *Computers & Security* 25 (4)**,** 289-296.

Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor, L., Hong, J., Blair, M. A. and Pham, T. (2009). "School of Phish: A Real-World Evaluation of Anti-Phishing Training", *5th Symposium on Usable Privacy and Security*, Mountain View, CA, USA.

Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J. and Nunge, E. (2007). "Protecting People from Phishing: The Design and Evaluation of an Embed Ded Training Email System", *SIGCHI Conference on Human factors in Computing Systems*, San José, CA, USA.

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F. and Hong, J. (2008). "Lessons from a Real World Evaluation of Anti-Phishing Training", *2008 eCrime Researchers Summit*, Atlanta, USA.

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F. and Hong, J. (2010). "Teaching Johnny Not to Fall for Phish", *ACM Transactions on Internet Technology (TOIT)* 10 (2)**,** 1-31.

Kwak, Y., Lee, S., Damiano, A. and Vishwanath, A. (2020). "Why Do Users Not Report Spear Phishing Emails?", *Telematics and Informatics* 48**,** 101343.

Lai, I. K. W. and Wong, J. W. C. (2020). "Comparing Crisis Management Practices in the Hotel Industry between Initial and Pandemic Stages of Covid-19", *International Journal of Contemporary Hospitality Management* 32 (10).

Lastdrager, E., Gallardo, I. C., Hartel, P. and Junger, M. (2017). "How Effective Is Anti-Phishing Training for Children?", *13th Symposium on usable Privacy and Security*, Santa Clara, CA, USA.

Lebek, B., Uffen, J., Neumann, M., Hohler, B. and Breitner, M. H. (2014). "Information Security Awareness and Behavior: A Theory-Based Literature Review", *Management Research Review* 37 (12).

Lee, A. S., Manoj, T. and Baskerville, R. L. (2015). "Going Back to Basics in Design Science: From the Information Technology Artifact to the Information Systems Artifact", *Information Systems Journal* 25 (1)**,** 5-21.

Lee, H. C. B., Cruz, J. M. and Shankar, R. (2018). "Corporate Social Responsibility (Csr) Issues in Supply Chain Competition: Should Greenwashing Be Regulated?", *Decision Sciences* 49 (6)**,** 1088-1115.

Lee, L. and Charles, V. (2021). "The Impact of Consumers' Perceptions Regarding the Ethics of Online Retailers and Promotional Strategy on Their Repurchase Intention", *International Journal of Information Management* 57**,** 102264.

Li, H., Luo, X. R. and Chen, Y. (2021). "Understanding Information Security Policy Violation from a Situational Action Perspective", *Journal of the Association for Information Systems* 22 (3)**,** 5.

Li, H., Yu, L. and He, W. (2019). "The Impact of GDPR on Global Technology Development", *Journal of Global Information Technology Management* 22 (1)**,** 1-6.

Li, L., Helenius, M. and Berki, E. (2012). "A Usability Test of Whitelist and Blacklist-Based Anti-Phishing Applic Ation", *16th International Academic MindTrek Conference*, Tampere, Finland.

Lin, E., Greenberg, S., Trotter, E., Ma, D. and Aycock, J. (2011). "Does Domain Highlighting Help People Identify Phishing Sites?", *SIGCHI Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada.

Linden, T., Khandelwal, R., Harkous, H. and Fawaz, K. (2020). "The Privacy Policy Landscape after the GDPR", *Proceedings on Privacy Enhancing Technologies* (1)**,** 47-64.

Linsner, S., Steinbrink, E., Kuntke, F., Franken, J. and Reuter, C. (2022). "Supporting Users in Data Disclosure Scenarios in Agriculture through Transparency", *Behaviour & Information Technology* 41 (10)**,** 2151-2173.

Liu, B. L., Pavlou, P. A. and Cheng, X. F. (2021). "Achieving a Balance between Privacy Protection and Data Collection: A Field Experimental Examination of a Theory-Driven Information Technology Solution", *Information Systems Research* 33 (1)**,** 203-223.

Liu, Z., Wang, X., Min, Q. and Li, W. (2019). "The Effect of Role Conflict on Self-Disclosure in Social Network Sites: An Integrated Perspective of Boundary Regulation and Dual Process Model", *Information Systems Journal* 29 (2)**,** 279-316.

Lowry, P. B., Dinev, T. and Willison, R. (2017). "Why Security and Privacy Research Lies at the Centre of the Information Systems (IS) Artefact: Proposing a Bold Research Agenda", *European Journal of Information Systems* 26 (6)**,** 546-563.

Lowry, P. B., Zhang, J., Moody, G. D., Chatterjee, S., Wang, C. and Wu, T. (2019). "An Integrative Theory Addressing Cyberharassment in the Light of Technology-Based Opportunism", *Journal of Management Information Systems* 36 (4)**,** 1142-1178.

Luo, X., Zhang, W., Burd, S. and Seazzu, A. (2013). "Investigating Phishing Victimization with the Heuristic–Systematic Model: A Theoretical Framework and an Exploration", *Computers & Security* 38**,** 28-38.

MacKenzie, S. B., Podsakoff, P. M. and Podsakoff, N. P. (2011). "Construct Measurement and Validation Procedures in Mis and Behavioral Research: Integrating New and Existing Techniques", *MIS Quarterly***,** 293-334.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S. and Söllner, M. (2019). "Ai-Based Digital Assistants", *Business & Information Systems Engineering* 61 (4)**,** 535-544.

Marforio, C., Jayaram Masti, R., Soriente, C., Kostiainen, K. and Čapkun, S. (2016). "Evaluation of Personalized Security Indicators as an Anti-Phishing Mec Hanism for

Smartphone Applications", *CHI Conference on Human Factors in Computing Systems*, San José, USA.

Marsch, M., Grossklags, J. and Patil, S. (2021). "Won't You Think of Others?: Interdependent Privacy in Smartphone App Permissions", *Proceedings of the ACM on Human-Computer Interaction* 5 (CSCW2)**,** 1-35.

Marsden, J., Albrecht, Z., Berggren, P., Halbert, J., Lemons, K., Moncivais, A. and Thompson, M. (2020). "Facts and Stories in Phishing Training: A Replication and Extension",

Martin, K. (2016). "Understanding Privacy Online: Development of a Social Contract Approach to Privacy", *Journal of business ethics* 137 (3)**,** 551-569.

Mayring, P. (2004). "Qualitative Content Analysis", *A companion to qualitative research* 1 (2)**,** 159-176.

Mayring, P. (2014). *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution,* Klagenfurt.

McInnes, M. D. F., Moher, D., Thombs, B. D., McGrath, T. A., Bossuyt, P. M., Clifford, T., Cohen, J. F., Deeks, J. J., Gatsonis, C., Hooft, L. and others (2018). "Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies: The Prisma-Dta Statement", *Jama* 319 (4)**,** 388-396.

Mehrabian, A. and Russell, J. A. (1974). *An Approach to Environmental Psychology,* Cambridge, MA, US: The MIT Press.

Mirtsch, M., Pohlisch, J. and Blind, K. (2020). "International Diffusion of the Information Security Management System Standard Iso/Iec 27001: Exploring the Role of Culture", *European Conference On Information Systems*, Virtual conference.

Mithani, M. A. (2020). "Adaptation in the Face of the New Normal", *Academy of Management Perspectives* 34 (4)**,** 508-530.

Mitnick, K. D. and Simon, W. L. (2003). *The Art of Deception: Controlling the Human Element of Security*: John Wiley & Sons.

Miyamoto, D., Iimura, T., Blanc, G., Tazaki, H. and Kadobayashi, Y. (2014). "Eyebit: Eye-Tracking Approach for Enforcing Phishing Prevention Habits", *3rd International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS)*, Wroclaw, Poland.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. and Group, P. (2009). "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The Prisma Statement", *PLoS med* 6 (7).

Moody, G. D., Siponen, M. and Pahnila, S. (2018). "Toward a Unified Model of Information Security Policy Compliance", *MIS Quarterly* 42 (1).

Moore, G. C. and Benbasat, I. (1991). "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation", *Information Systems Research* 2 (3)**,** 192-222.

Naidoo, R. (2020). "A Multi-Level Influence Model to Covid-19 Themed Cybercrime", *European Journal of Information Systems* 29 (3)**,** 306-321.

NCATS. (2018). *United States Department of Homeland Security (Dhs) National Cybersecurity Assessments and Technical Services Ncats: Phishing Campaign Assessment Summary*. [Online] Available: https://www.cisa.gov/uscert/sites/default/files/resources/ncats/PCA%20Sample%20Report_508-Compliant.pdf (Accessed April 22, 2022).

Neuendorf, K. A. (2011). *Content Analysis Guidebook*: Sage Publications Incorporated.

Neumann, L. (2019). *Hirne Hacken - Menschliche Faktoren Der It Sicherheit, Vortrag Auf Dem 6. Chaos Communication Congress (36c3)*. [Online] Available: https://media.ccc.de/v/36c3-11175-hirne_hacken (Accessed January 13, 2020).

Nicholson, J., Coventry, L. and Briggs, P. (2017). "Can We Fight Social Engineering Attacks by Social Means? Assessing Social Salience as a Means to Improve Phish Detection", *13th Symposium on Usable Privacy and Security,* Santa Clara, CA, USA.

Nissenbaum, H. (2004). "Privacy as Contextual Integrity", *Wash. L. Rev.* 79**,** 119.

NIST. (2011). *Managing Information Security Risk: Organization, Mission, and Information System View*. National Institute of Standards and Technology, U.S. Department of Commerce. [Online] Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-39.pdf (Accessed January 14, 2023).

Nouwens, M., Liccardi, I., Veale, M., Karger, D. and Kagal, L. (2020). "Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence", *2020 CHI Conference on Human Factors in Computing Systems,* Virtual conference.

Nunamaker, J., Jay F and Briggs, R. O. (2012). "Toward a Broader Vision for Information Systems", *ACM Transactions on Management Information Systems (TMIS)* 2 (4)**,** 1-12.

O'Shaughnessy, J. and O'Shaughnessy, N. J. (2002). "Marketing, the Consumer Society and Hedonism", *European Journal of Marketing* 36 (5)**,** 524-547.

Oh, H. J., Kim, J., Chang, J. J., Park, N. and Lee, S. (2023). "Social Benefits of Living in the Metaverse: The Relationships among Social Presence, Supportive Interaction, Social Self-Efficacy, and Feelings of Loneliness", *Computers in Human Behavior* 139**,** 107498.

Olson, I. R., Plotzker, A. and Ezzyat, Y. (2007). "The Enigmatic Temporal Pole: A Review of Findings on Social and Emotional Processing", *Brain* 130 (7)**,** 1718-1731.

Olteanu, A.-M., Huguenin, K., Dacosta, I. and Hubaux, J.-P. (2018). "Consensual and Privacy-Preserving Sharing of Multi-Subject and Interdependent Data", *25th Network and Distributed System Security Symposium (NDSS),* San Diego, CA, USA.

Palan, S. and Schitter, C. (2018). "Prolific.Ac-a Subject Pool for Online Experiments", *Journal of Behavioral and Experimental Finance* 17**,** 22-27.

Patsakis, C. and Chrysanthou, A. (2020). “Analysing the Fall 2020 Emotet Campaign”, *arXiv preprint arXiv:2011.06479*.

Pavlou, P. A. (2011). “State of the Information Privacy Literature: Where Are We Now and Where Should We Go?”, *MIS Quarterly***,** 977-988.

Pavlou, P. A., Liang, H. G. and Xue, Y. J. (2007). “Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective”, *MIS Quarterly* 31 (1)**,** 105-136.

Peer, E., Egelman, S., Harbach, M., Malkin, N., Mathur, A. and Frik, A. (2020). “Nudge Me Right: Personalizing Online Nudges to People's Decision-Making Styles”, *Computers in Human Behavior* 109.

Pereira, M. M. O., Silva, M. E. and Hendry, L. C. (2021). “Supply Chain Sustainability Learning: The Covid-19 Impact on Emerging Economy Suppliers”, *Supply Chain Management: An International Journal* 26 (6).

Perrault, E. K. (2018). “Using an Interactive Online Quiz to Recalibrate College Students’ Atti Tudes and Behavioral Intentions About Phishing”, *Journal of Educational Computing Research* 55 (8)**,** 1154-1167.

Petelka, J., Zou, Y. and Schaub, F. (2019). “Put Your Warning Where Your Link Is: Improving and Evaluating Email Phishing Warnings”, *CHI Conference on Human Factors in Computing Systems - CH '19*, Glasgow, UK.

Petronio, S., Child, J. T. and Hall, R. D. (2021). “Communication Privacy Management Theory: Significance for Interpersonal Communication”, (eds.) *Engaging Theories in Interpersonal Communication*, 314-327. Routledge.

Piccoli, G. (2016). “Triggered Essential Reviewing: The Effect of Technology Affordances on Service Experience Evaluations”, *European Journal of Information Systems* 25 (6)**,** 477-492.

Pienta, D., Thatcher, J. B. and Johnston, A. (2020). “Protecting a Whale in a Sea of Phish”, *Journal of Information Technology* 35 (3)**,** 214-231.

Pienta, D., Thatcher, J. B. and Johnston, A. C. (2018). “A Taxonomy of Phishing: Attack Types Spanning Economic, Temporal, Breadth, and Target Boundaries”, *WISP 2018*, San Francisco, CA, USA.

Pirlott, A. G. and MacKinnon, D. P. (2016). “Design Approaches to Experimental Mediation”, *Journal of experimental social psychology* 66**,** 29-38.

Posey, C., Roberts, T. L. and Lowry, P. B. (2015). “The Impact of Organizational Commitment on Insiders’ Motivation to Protect Organizational Information Assets”, *Journal of Management Information Systems* 32 (4)**,** 179-214.

Posey, C., Roberts, T. L., Lowry, P. B. and Hightower, R. T. (2014). “Bridging the Divide: A Qualitative Comparison of Information Security Thought Patterns between Information Security Professionals and Ordinary Organizational Insiders”, *Information & Management* 51 (5)**,** 551-567.

Pradhan, A., Findlater, L. and Lazar, A. (2019). “"Phantom Friend" or" Just a Box with Information": Personification and Ontological Categorization of Smart Speaker-Based Voice Assistants by Older Adults”, *Proceedings of the ACM on Human-Computer Interaction* 3 (CSCW)**,** 1-21.

Pradies, C., Aust, I., Bednarek, R., Brandl, J., Carmine, S., Cheal, J., Pina e Cunha, M., Gaim, M., Keegan, A. and Lê, J. K. (2021). “The Lived Experience of Paradox: How Individuals Navigate Tensions During the Pandemic Crisis”, *Journal of Management Inquiry* 30 (2)**,** 154-167.

Primack, B. A., Sidani, J., Agarwal, A. A., Shadel, W. G., Donny, E. C. and Eissenberg, T. E. (2008). “Prevalence of and Associations with Waterpipe Tobacco Smoking among Us University Students”, *Annals of Behavioral Medicine* 36 (1)**,** 81-86.

Proofpoint. (2019). *The Human Factor Report*. [Online] Available: https://www.proofpoint.com/us/resources/threat-reports/human-factor (Accessed November 15, 2020).

Pu, Y. and Grossklags, J. (2015). “Using Conjoint Analysis to Investigate the Value of Interdependent Privacy in Social App Adoption Scenarios”, *Proceedings of the International Conference on Information Systems*, Fort Worth, United States.

Pu, Y. and Grossklags, J. (2016). “Towards a Model on the Factors Influencing Social App Users' Valuation of Interdependent Privacy”, *Proc. Priv. Enhancing Technol.* 2016 (2)**,** 61-81.

Pu, Y. and Grossklags, J. (2017). “Valuating Friends’ Privacy: Does Anonymity of Sharing Personal Data Matter?”, *13th Symposium on Usable Privacy and Security*, Santa Clara, CA, USA.

PwC. (2022). *Pwc's 25th Global CEO Survey on Cybersecurity*. [Online] Available: https://www.pwc.de/de/ceosurvey/2022/pwc-25th-ceo-survey.pdf (Accessed January 5, 2022).

Reeder, R. W., Felt, A. P., Consolvo, S., Malkin, N., Thompson, C. and Egelman, S. (2018). “An Experience Sampling Study of User Reactions to Browser Warnings in the Field”, *Proceedings of the 2018 CHI conference on human factors in computing systems*, Montréal, Canada.

Reinheimer, B., Aldag, L., Mayer, P., Mossano, M., Duezguen, R., Lofthouse, B., Landesberger, T. and Volkamer, M. (2020). “An Investigation of Phishing Awareness and Education over Time: When and How to Best Remind Users”, *16th Symposium on Usable Privacy and Security*, Boston, MA, USA.

Renaud, K. and Dupuis, M. (2019). “Cyber Security Fear Appeals: Unexpectedly Complicated”, *New Security Paradigms Workshop*, San Carlos, Cost Rica.

Renaud, K. and Zimmermann, V. (2019). “Nudging Folks Towards Stronger Password Choices: Providing Certainty Is the Key”, *Behavioural Public Policy* 3 (2)**,** 228-258.

Renaud, K., Zimmermann, V., Schürmann, T. and Böhm, C. (2021). "Exploring Cybersecurity-Related Emotions and Finding That They Are Challenging to Measure", *Humanities and Social Sciences Communications* 8 (1)**,** 1-17.

Reuter, C., Iacono, L. L. and Benlian, A. (2022). "A Quarter Century of Usable Security and Privacy Research: Transparency, Tailorability, and the Road Ahead", *Behaviour & Information Technology* 41 (10)**,** 1-14.

Richards, L. (2014). *Handling Qualitative Data: A Practical Guide*: Sage.

Roethke, K., Klumpe, J., Adam, M. and Benlian, A. (2020). "Social Influence Tactics in E-Commerce Onboarding: The Role of Social Proof and Reciprocity in Affecting User Registrations", *Decision Support Systems* 131**,** 113268.

Romanelli, E. and Tushman, M. L. (1994). "Organizational Transformation as Punctuated Equilibrium: An Empirical Test", *Academy of Management Journal* 37 (5)**,** 1141-1166.

Ronda, T., Saroiu, S. and Wolman, A. (2008). "Itrustpage: A User-Assisted Anti-Phishing Tool", *ACM SIGOPS Operating Systems Review* 42 (4)**,** 261-272.

Ruighaver, A. B., Maynard, S. B. and Chang, S. (2007). "Organisational Security Culture: Extending the End-User Perspective", *Computers & Security* 26 (1)**,** 56-62.

Samonas, S. and Coss, D. (2014). "The Cia Strikes Back: Redefining Confidentiality, Integrity and Availability in Security", *Journal of Information System Security* 10 (3).

Sarker, S., Chatterjee, S., Xiao, X. and Elbanna, A. (2019). "The Sociotechnical Axis of Cohesion for the IS Discipline: Its Historical Legacy and Its Continued Relevance", *MIS Quarterly* 43 (3)**,** 659-719.

Sasse, A. (2015). "Scaring and Bullying People into Security Won't Work", *IEEE Security & Privacy* 13 (3)**,** 80-83.

Schechter, S. E., Dhamija, R., Ozment, A. and Fischer, I. (2007). "The Emperor's New Security Indicators", *IEEE Symposium on Security and Privacy*, Berkeley, CA, USA.

Schein, E. H. (2010). *Organizational Culture and Leadership*, 2. John Wiley & Sons.

Schlienger, T. and Teufel, S. (2003a). "Analyzing Information Security Culture: Increased Trust by an Appropriate Information Security Culture", *14th International Workshop on Database and Expert Systems Applications*, Prague, Czech Republic.

Schlienger, T. and Teufel, S. (2003b). "Information Security Culture: From Analysis to Change", *Annual Information Security South Africa Conference*, Johannesburg, South Africa.

Schneider, C., Weinmann, M. and Vom Brocke, J. (2018). "Digital Nudging: Guiding Online User Choices through Interface Design", *Communications of the ACM* 61 (7)**,** 67-73.

Schöbel, S., Barev, T., Janson, A., Hupfeld, F. and Leimeister, J. M. (2020). "Understanding User Preferences of Digital Privacy Nudges–a Best-Worst Scaling Approach", *53rd Hawaii International Conference on System Sciences*, Maui, HI, USA.

Schryen, G., Wagner, G., Benlian, A. and Paré, G. (2020). "A Knowledge Development Perspective on Literature Reviews: Validation of a New Typology in the Is Field", *Communications of the AIS* 46 (7).

Schuetz, S. W., Lowry, P. B., Pienta, D. A. and Thatcher, J. B. (2020). "The Effectiveness of Abstract Versus Concrete Fear Appeals in Information Security", *Journal of Management Information Systems* 37 (3)**,** 723-757.

Scott, M. J., Ghinea, G. and Arachchilage, N. A. G. (2014). "Assessing the Role of Conceptual Knowledge in an Anti-Phishing Educational Game", *14th International Conference on Advanced Learning Technologies*, Athens, Greece.

Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L. F., Hong, J. and Nunge, E. (2007). "Anti-Phishing Phil: The Design and Evaluation of a Game That Teaches People Not to Fall for Phish", *3rd Symposium on Usable Privacy and Security*, Pittsburgh, PA, USA.

Silic, M. and Lowry, P. B. (2020). "Using Design-Science Based Gamification to Improve Organizational Security Training and Compliance", *Journal of Management Information Systems* 37 (1)**,** 129-161.

Silva, L. and Hirschheim, R. (2007). "Fighting against Windmills: Strategic Information Systems and Organizational Deep Structures", *MIS Quarterly* 31 (2)**,** 327-354.

Singer, T. and Ricard, M. (2015). *Caring Economics: Conversations on Altruism and Compassion, between Scientists, Economists, and the Dalai Lama*: Picador.

Siponen, M. T. (2000). "A Conceptual Foundation for Organizational Information Security Awareness", *Information Management & Computer Security* 8 (1).

Siponen, M. T. (2001). "Five Dimensions of Information Security Awareness", *SIGCAS Comput. Soc.* 31 (2)**,** 24-29.

Smith, H. J., Milberg, S. J. and Burke, S. J. (1996). "Information Privacy: Measuring Individuals' Concerns About Organizational Practices", *MIS Quarterly***,** 167-196.

Solove, D. J. (2008). *Understanding Privacy*, Cambridge, MA: Harvard University Press.

Sorensen, M. (2018). *The New Face of Phishing*. https://apwg.org/the-new-face-of-phishing/. [Online] Available: https://apwg.org/the-new-face-of-phishing/ (Accessed January 13, 2021).

Spiekermann, S., Krasnova, H., Hinz, O., Baumann, A., Benlian, A., Gimpel, H., Heimbach, I., Köster, A., Maedche, A. and Niehaves, B. (2022). "Values and Ethics in Information Systems", *Business & Information Systems Engineering* 64 (2)**,** 247-264.

Squicciarini, A. C., Shehab, M. and Paci, F. (2009). "Collective Privacy Management in Social Networks", *18th International Conference on World Wide Web*, Madrid, Spain.

Statista. (2021). *Most Popular Social Networks Worldwide as of July 2021, Ranked by Number of Active Users*. [Online] Available: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ (Accessed August 04, 2021).

Stembert, N., Padmos, A., Bargh, M. S., Choenni, S. and Jansen, F. (2015). "A Study of Preventing Email (Spear) Phishing by Enabling Human Intelli Gence", *European Intelligence and Security Informatics Conference*, Manchester, UK.

Stockhardt, S., Reinheimer, B., Volkamer, M., Mayer, P., Kunz, A., Rack, P. and Lehmann, D. (2016). "Teaching Phishing-Security: Which Way Is Best?", *IFIP International Conference on ICT Systems Security and Privacy Protection*, Ghent, Belgium.

Sutanto, J., Tan, E. P. C.-H. and Phang, C. W. (2013). "Addressing the Personalization-Privacy Paradox: An Empirical Assessment from a Field Experiment on Smartphone Users", *MIS Quarterly* 37 (4).

Sutanto, J., Wenninger, H. and Duriana, H. (2021). "Warm-Glow Giving, Hedonism, and Their Influence on Muslim User Engagement on Loan-Based Crowdfunding Platforms", *Journal of the Association for Information Systems* 22 (2)**,** 7.

Symantec. (2019). *Internet Security Threat Report, Volume 24*. [Online] Available: https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/istr-24-cyber-security-threat-landscape (Accessed January 10, 2022).

Symeonidis, I., Biczók, G., Shirazi, F., Pérez-Solà, C., Schroers, J. and Preneel, B. (2018). "Collateral Damage of Facebook Third-Party Applications: A Comprehensive Study", *Computers & Security* 77**,** 179-208.

Szilagyi, A. D. and Wallace, M. J. (1983). *Organizational Behavior and Performance*: Good Year Books.

Taylor, S. and Todd, P. A. (1995). "Understanding Information Technology Usage: A Test of Competing Models", *Information Systems Research* 6 (2)**,** 144-176.

Thaler, R. and Sunstein, C. (2008). *Nudge: Improving Decisions About Health, Wealth and Happiness Penguin*: Penguin Books, New York.

The Guardian. (2020). *'Shocking' Hack of Psychotherapy Records in Finland Affects Thousands*. [Online] Available: https://www.theguardian.com/world/2020/oct/26/tens-of-thousands-psychotherapy-records-hacked-in-finland (Accessed January 09, 2023).

The Radicati Group Inc. . (2019). *Email Statistics Report, 2019-2023 Executive Summary*. https://www.radicati.com/?download=email-statistics-report-2019-2023. [Online] Available: https://www.radicati.com/?download=email-statistics-report-2019-2023 (Accessed February 8, 2022).

Thomas, K., Grier, C. and Nicol, D. M. (2010). "Unfriendly: Multi-Party Privacy Risks in Social Networks", *10th Privacy Enhancing Technologies Symposium*, Berlin, Deutschland.

Tian, K., Jan, S. T. K., Hu, H., Yao, D. and Wang, G. (2018). "Needle in a Haystack: Tracking Down Elite Phishing Domains in the Wild",

Tidy, J. (2021). *Colonial Hack: How Did Cyber-Attackers Shut Off Pipeline?* [Online] Available: https://www.bbc.com/news/technology-57063636 (Accessed January 08, 2023).

Tim, Y., Pan, S. L., Bahri, S. and Fauzi, A. (2018). "Digitally Enabled Affordances for Community-Driven Environmental Movement in Rural Malaysia", *Information Systems Journal* 28 (1)**,** 48-75.

Tsai, J. Y., Egelman, S., Cranor, L. and Acquisti, A. (2011). "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study", *Information Systems Research* 22 (2)**,** 254-268.

Tsohou, A., Karyda, M., Kokolakis, S. and Kiountouzis, E. (2015). "Managing the Introduction of Information Security Awareness Programmes in Organisations", *European Journal of Information Systems* 24 (1)**,** 38-58.

Turel, O., He, Q. and Wen, Y. (2021). "Examining the Neural Basis of Information Security Policy Violations: A Noninvasive Brain Stimulation Approach", *MIS Quarterly* 45 (4)**,** 1715-1744.

Turel, O., Matt, C., Trenz, M. and Cheung, C. M. (2020). "An Intertwined Perspective on Technology and Digitised Individuals: Linkages, Needs and Outcomes", *Information Systems Journal* 30 (6)**,** 929-939.

Tushman, M. and Romanelli, E. (1985). "Convergence and Reorientation: A Metamorphosis Model", *Research in Organizational Behavior* 7**,** 171-222.

Valentine, K. A., Li, N. P., Penke, L. and Perrett, D. I. (2014). "Judging a Man by the Width of His Face: The Role of Facial Ratios and Dominance in Mate Choice at Speed-Dating Events", *Psychological science* 25 (3)**,** 806-811.

Van der Heijden, H. (2004). "User Acceptance of Hedonic Information Systems", *MIS Quarterly* 28 (4)**,** 695-704.

Van Niekerk, J. and Von Solms, R. (2005). "A Holistic Framework for the Fostering of an Information Security Sub-Culture in Organizations", *Annual Information Security South Africa Conference*, Johannesburg, South Africa.

Van Niekerk, J. F. and Von Solms, R. (2010). "Information Security Culture: A Management Perspective", *Computers & Security* 29 (4)**,** 476-486.

Van Schaik, P., Jeske, D., Onibokun, J., Coventry, L., Jansen, J. and Kusev, P. (2017). "Risk Perceptions of Cyber-Security and Precautionary Behaviour", *Computers in Human Behavior* 75**,** 547-559.

Vance, A., Jenkins, J. L., Anderson, B. B., Bjornn, D. K. and Kirwan, C. B. (2018). "Tuning out Security Warnings: A Longitudinal Examination of Habituation through Fmri, Eye Tracking, and Field Experiments", *MIS Quarterly* 42 (2)**,** 355-380.

Varshney, G., Sardana, A. and Joshi, R. C. (2012). "Secret Information Display Based Authentication Technique Towards Prev Enting Phishing Attacks", *International Conference on Advances in Computing, Communications and Informatics*, Madras, India.

Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D. (2003). "User Acceptance of Information Technology: Toward a Unified View", *MIS Quarterly* 27 (3)**,** 425-478.

Venkatesh, V., Thong, J. Y. and Xu, X. (2012). "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology", *MIS Quarterly* 36 (1)**,** 157-178.

Verizon. (2022). *Data Breach Investigations Report*. [Online] Available: https://www.verizon.com/business/resources/T871/reports/dbir/2022-data-breach-investigations-report-dbir.pdf (Accessed January 12, 2023).

Verma, R. and Dyer, K. (2015). "On the Character of Phishing Urls: Accurate and Robust Statistical Lea Rning Classifiers", *5th ACM Conference on Data and Application Security and Privacy*, San Antonio, TX, USA.

Vielberth, M., Englbrecht, L. and Pernul, G. (2021). "Improving Data Quality for Human-as-a-Security-Sensor. A Process Driven Quality Improvement Approach for User-Provided Incident Information", *Information & Computer Security* 29 (2).

Volkamer, M., Renaud, K., Reinheimer, B. and Kunz, A. (2017). "User Experiences of Torpedo: Tooltip-Powered Phishing Email Detection", *Computers & Security* 71**,** 100-113.

Von Solms, B. and Von Solms, R. (2005). "From Information Security to… Business Security?", *Computers & Security* 24 (4)**,** 271-273.

Vroom, C. and von Solms, R. (2004). "Towards Information Security Behavioural Compliance", *Computers & Security* 23 (3)**,** 191-198.

Waizenegger, L., McKenna, B., Cai, W. and Bendz, T. (2020). "An Affordance Perspective of Team Collaboration and Enforced Working from Home During Covid-19", *European Journal of Information Systems* 29 (4)**,** 429-442.

Wakefield, R. L. and Whitten, D. (2006). "Mobile Computing: A User Study on Hedonic/Utilitarian Mobile Device Usage", *European Journal of Information Systems* 15 (3)**,** 292-300.

Wang, J., Li, Y. and Rao, H. R. (2017). "Coping Responses in Phishing Detection: An Investigation of Antecedents and Consequences", *Information Systems Research* 28 (2)**,** 378-396.

Wang, Y., Leon, P. G., Acquisti, A., Cranor, L. F., Forget, A. and Sadeh, N. (2014). "A Field Trial of Privacy Nudges for Facebook", *SIGCHI Conference on Human Factors in Computing Systems*, Toronto, Canada.

Warkentin, M., Goel, S. and Menard, P. (2017). "Shared Benefits and Information Privacy: What Determines Smart Meter Technology Adoption?", *Journal of the Association for Information Systems* 18 (11).

Wash, R. (2020). "How Experts Detect Phishing Scam Emails", *Proceedings of the ACM on Human-Computer Interaction* 4 (CSCW2)**,** 1-28.

Wash, R. and Cooper, M. M. (2018). "Who Provides Phishing Training? Facts, Stories, and People Like Me", *CHI Conference on Human Factors in Computing Systems*, Montréal, Canada.

Wash, R., Nthala, N. and Rader, E. (2021). "Knowledge and Capabilities That Non-Expert Users Bring to Phishing Detection", *17th Symposium on Usable Privacy and Security,* Virtual.

Weanquoi, P., Johnson, J. and Zhang, J. (2017). "Using a Game to Teach About Phishing", *18th Annual Conference on Information Technology Education,* Rochester, NY, USA.

Weber, R. P. (1990). *Basic Content Analysis*: Sage.

Weinmann, M., Schneider, C. and Vom Brocke, J. (2016). "Digital Nudging", *Business & Information Systems Engineering* 58 (6)**,** 433-436.

Wen, Z. A., Lin, Z., Chen, R. and Andersen, E. (2019). "What. Hack: Engaging Anti-Phishing Training through a Role-Playing Phishing Simulation Game", *CHI Conference on Human Factors in Computing Systems,* Glasgow, UK.

Werner, D., Adam, M. and Benlian, A. (2022). "Empowering Users to Control Ads and Its Effects on Website Stickiness", *Electronic Markets* 32 (3)**,** 1373-1397.

Westin, A. F. (1970). *Privacy and Freedom,* New York, NY: Ig Publishing.

Widup, S., Spitler, M., Hylender, D. and Bassett, G. (2018). *2018 Verizon Data Breach Investigations Report*. [Online] Available: http://www.verizonenterprise.com/de/DBIR/ (Accessed July 02, 2020).

Wiese, O., Lausch, J., Bode, J. and Roth, V. (2018). "Beware the Downgrading of Secure Electronic Mail", *8th Workshop on Socio-Technical Aspects in Security and Trust,* San Juan, Puerto Rico, USA.

Wiley, A., McCormac, A. and Calic, D. (2020). "More Than the Individual: Examining the Relationship between Culture and Information Security Awareness", *Computers & Security* 88**,** 101640.

Williams, E. J., Hinds, J. and Joinson, A. N. (2018). "Exploring Susceptibility to Phishing in the Workplace", *International Journal of Human-Computer Studies* 120**,** 1-13.

Willison, R. and Warkentin, M. (2013). "Beyond Deterrence: An Expanded View of Employee Computer Abuse", *MIS Quarterly* 37 (1)**,** 1-20.

Wirth, J., Maier, C., Laumer, S. and Weitzel, T. (2019). "Perceived Information Sensitivity and Interdependent Privacy Protection: A Quantitative Study", *Electronic Markets* 29 (3)**,** 359-378.

Wixom, B. H. and Todd, P. A. (2005). "A Theoretical Integration of User Satisfaction and Technology Acceptance", *Information Systems Research* 16 (1)**,** 85-102.

Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M. and Marett, K. (2014a). "Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance", *Information Systems Research* 25 (2)**,** 385-400.

Wright, R. T. and Marett, K. (2010). "The Influence of Experiential and Dispositional Factors in Phishing: An Empirical Investigation of the Deceived", *Journal of Management Information Systems* 27 (1)**,** 273-303.

Wright, R. T., Marett, K. and Thatcher, J. B. (2014b). "Extending Ecommerce Deception to Phishing", *International Conference on Information Systems,* Auckland, New Zealand.

Wu, M., Miller, R. C. and Garfinkel, S. L. (2006a). "Do Security Toolbars Actually Prevent Phishing Attacks?", *SIGCHI conference on Human Factors in computing systems,* Montréal, Canada.

Wu, M., Miller, R. C. and Little, G. (2006b). "Web Wallet: Preventing Phishing Attacks by Revealing User Intentions", *2nd symposium on Usable Privacy and Security,* Pittsburgh, PA, USA.

Xiong, A., Proctor, R. W., Yang, W. and Li, N. (2019). "Embedding Training within Warnings Improves Skills of Identifying Phishing Webpages", *Human Factors* 61 (4)**,** 577-595.

Yang, W., Xiong, A., Chen, J., Proctor, R. W. and Li, N. (2017). "Use of Phishing Training to Improve Security Warning Compliance: Evidence from a Field Experiment", *Hot Topics in Science of Security: Symposium and Bootcamp,* Hanover, MD, USA.

Yao, H. and Shin, D. (2013). "Towards Preventing QR Code Based Attacks on Android Phone Using Security Warnings", *8th ACM SIGSAC Symposium on Information, Computer and Communications Security,* Hangzhou China.

Yee, K.-P. and Sitaker, K. (2006). "Passpet: Convenient Password Management and Phishing Protection", *2nd Symposium on Usable Privacy and Security* Pittsburgh, PA, USA.

Yin, R. K. (2017). *Case Study Research and Applications: Design and Methods*: Sage publications.

Yue, C. (2012). "Preventing the Revealing of Online Passwords to Inappropriate Websites with Logininspector", *26th Large Installation System Administration Conference,* San Diego, CA, USA.

Zhang-Kennedy, L. and Chiasson, S. (2021). "A Systematic Review of Multimedia Tools for Cybersecurity Awareness an D Education", *ACM Computing Surveys (CSUR)* 54 (1)**,** 1-39.

Zhang, B. and Xu, H. (2016). "Privacy Nudges for Mobile Applications: Effects on the Creepiness Emotion and Privacy Attitudes", *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing,* San Francisco, CA, USA.

Zhao, X., Lynch Jr, J. G. and Chen, Q. (2010). "Reconsidering Baron and Kenny: Myths and Truths About Mediation Analysis", *Journal of Consumer Research* 37 (2)**,** 197-206.

Zimmermann, V. and Renaud, K. (2019). "Moving from a 'Human-as-Problem" to a 'Human-as-Solution" Cybersecurity Mindset", *International Journal of Human-Computer Studies* 131**,** 169-187.

Zimmermann, V. and Renaud, K. (2021). “The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions”, *ACM Transactions on Computer-Human Interaction* 28 (1)**,** 7:1 - 7:45.

Declaration of Authorship

I hereby declare that the submitted dissertation is my own work. All quotes, whether word by word or in my own words, have been marked as such. The dissertation has not been published anywhere else nor presented to any other examination board.

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig angefertigt habe. Sämtliche aus fremden Quellen direkt oder indirekt übernommene Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder einer anderen Prüfungsbehörde vorgelegt noch veröffentlicht.

_____

Anjuli Franz

Darmstadt, 25.01.2023